الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الجزائر3-كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

-مطبوعة بيداغوجية

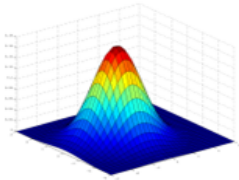
-موجهة لطلبة السنة الثانية علوم التسيير

-طور الليسانس

تحت عنوان

احصاء 3

مدعمة بتمارين



من اعداد الدكتور: بن الدين دوادي -أستاذ محاضر قسم أ

السنة الجامعية 2026/2025

بسم الله الرحمن الرحيم

فہرس

فهرس

فهرس	الصفحة
مقدمة	أ
المحاضرة الأولى: توزيع برنولي وتوزيع ذي الحدين (الثنائي)	**1**
1-توزيع برنولي	**3**
1-1-التعريف	**3**
1-2-المميزات العددية للتوزيع	**4**
1-3-التمثيل البياني للتوزيع	**4**
2-توزيع ذي الحدين (الثنائي)	**6**
2-1-التعريف	**6**
2-2-المميزات العددية للتوزيع	**6**
2-3-التمثيل البياني للتوزيع	**7**
3-تمارين محلولة	**8**
المحاضرة الثانية: توزيع بواسون	**13**
1-التعريف	**16**
2-المميزات العددية للتوزيع	**17**
3-التمثيل البياني للتوزيع	**17**
4-تمارين محلولة	**18**
المحاضرة الثالثة: التوزيع الهندسي	**24**
1-التعريف	**27**
2-المميزات العددية للتوزيع	**27**
3-التمثيل البياني للتوزيع	**28**

4-تمارين محلولة	**29**
المحاضرة الرابعة: التوزيع فوق الهندسي	**35**
1-التعريف	**38**
2-المميزات العددية للتوزيع	**38**
3-التمثيل البياني للتوزيع	**39**
4-تمارين محلولة	**40**
المحاضرة الخامسة: التوزيع الطبيعي	**46**
1-التعريف	**49**
2-المميزات العددية للتوزيع	**50**
3-التمثيل البياني للتوزيع	**50**
4-التوزيع الطبيعي القياسي	**51**
5-الانتقال من التوزيع الطبيعي الى التوزيع الطبيعي القياسي	**51**
6-تمارين محلولة	**52**
المحاضرة السادسة: التوزيع المنتظم	**58**
1-التعريف	**61**
2-المميزات العددية للتوزيع	**63**
3-التمثيل البياني للتوزيع	**64**
4-تمارين محلولة	**65**
المحاضرة السابعة: التوزيع الأسي	**70**
1-التعريف	**73**
2-المميزات العددية للتوزيع	**74**
3-التمثيل البياني للتوزيع	**74**
4-تمارين محلولة	**75**
المحاضرة الثامنة: التوزيع قاما	**79**
1-التعريف	**82**
2-المميزات العددية للتوزيع	**83**
3-التمثيل البياني للتوزيع	**83**

4-تمارين محلولة	**84**
المحاضرة التاسعة: التوزيع بيتا	**89**
1-التعريف	**92**
2-المميزات العددية للتوزيع	**93**
3-التمثيل البياني للتوزيع	**93**
4-تمارين محلولة	**94**
المحاضرة العاشرة: التوزيع كاي مربع	**99**
1-التعريف	**102**
2-المميزات العددية للتوزيع	**103**
3-التمثيل البياني للتوزيع	**104**
4-تمارين محلولة	**105**
المحاضرة الحادية عشر: توزيع ستيودنت وتوزيع فيشر	**109**
1-التوزيع ستيودنت	**111**
1-1-التعريف	**112**
1-2-المميزات العددية للتوزيع	**113**
1-3-التمثيل البياني للتوزيع	**113**
2-التوزيع فيشر	**114**
2-1-التعريف	**115**
2-2-المميزات العددية للتوزيع	**116**
2-3-التمثيل البياني للتوزيع	**116**
3-تمارين محلولة	**117**
المحاضرة الثانية عشر: تقارب بعض التوزيعات الاحتمالية	**122**
1-تقارب التوزيعات الاحتمالية	**124**
2-أنواع التقارب الاحتمالي	**125**
2-1-التقارب في التوزيع	**125**
2-2-التقارب في الاحتمال	**125**
3-نظرية النهاية المركزية	**126**

4-استمرارية تقارب التوزيعات الاحتمالية	**127**
5-بعض أشهر تقارب التوزيعات الاحتمالية	**129**
5-1-تقارب التوزيع ذو الحدين بالتوزيع الطبيعي	**129**
5-2-تقارب توزيع بواسون بالتوزيع الطبيعي	**130**
5-3-تقارب توزيع ذو الحدين بتوزيع بواسون	**131**
5-4-تقارب توزيع كاي مربع بالتوزيع الطبيعي	**131**
6-تمارين محلولة	**131**
المحاضرة الثالثة عشر: المتغيرات العشوائية المزدوجة	**137**
1-المتغيرات العشوائية المزدوجة	**139**
2-المتغيرات العشوائية المزدوجة المنفصلة	**140**
3-المتغيرات العشوائية المزدوجة المتصلة	**141**
4-المميزات العددية للمتغيرات العشوائية المزدوجة	**142**
6-تمارين محلولة	**144**
المحاضرة الرابعة عشر: دوال المتغيرات العشوائية المزدوجة	**150**
1-دوال المتغيرات العشوائية المزدوجة	**152**
2-دوال المتغيرات العشوائية المزدوجة المنفصلة	**152**
3-دوال المتغيرات العشوائية المزدوجة المتصلة	**153**
4-التمثيل البياني لدوال المتغيرات العشوائية المزدوجة	**154**
4-1-التمثيل البياني لدوال المتغيرات المزدوجة المنفصلة	**154**
4-1-التمثيل البياني لدوال المتغيرات المزدوجة المتصلة	**155**
5-تمارين محلولة	**156**
قائمة المراجع	**159**

مقدمة

لقد تزايدت أهمية علم الاحصاء بتزايد تطبيقاته في مختلف المجالات العلمية، فلقد أصبح يستعان به من اجل معالجة بعض المشاكل التي تواجه العلوم الأخرى، وتزايدت استعمالاته في مجالات الاقتصاد الحديث لما يوفره من أدوات تساعد الاقتصاد لاتخاذ القرارات السليمة في الوقت والمكان المناسبين، ومع تطور نظرية الاحتمالات تطورت معه أدوات الإحصاء واستعمالاته التي تترافق مع النمذجة والتنبؤ والمحاكاة التي تحتاج الى حساب الاحتمالات.

نظرا لأهمية علم الإحصاء في الميدان الاقتصادي، أتت هذه المطبوعة لتتماشى مع المقرر الوزاري الذي اعتمدته الجامعات في تدريس مقياس الإحصاء 3، لطلبة العلوم الاقتصادية والتجارية وعلوم التسيير، من خلال أربعة محاور:

-المحور الأول: اهم قوانين التوزيعات الاحتمالية المتقطعة؛

-المحور الثاني: اهم قوانين التوزيعات الاحتمالية المتصلة؛

-المحور الثالث: تقارب التوزيعات الاحتمالية؛

-المحور الرابع: المتغيرات العشوائية الثنائية.

وأربعة عشر محاضرة:

المحاضرة الاولى	توزيع برنولي وتوزيع ذي الحدين (الثنائي)	المحاضرة الثامنة	التوزيع قاما
المحاضرة الثانية	توزيع بواسون	المحاضرة التاسعة	التوزيع بيتا
المحاضرة الثالثة	التوزيع الهندسي	المحاضرة العاشرة	التوزيع كاي مربع
المحاضرة الرابعة	التوزيع فوق الهندسي	المحاضرة الحادية عشر	توزيع ستيودنت وتوزيع فيشر
المحاضرة الخامسة	التوزيع الطبيعي	المحاضرة الثانية عشر	تقارب بعض التوزيعات الاحتمالية
المحاضرة السادسة	التوزيع المنتظم	المحاضرة الثالثة عشر	المتغيرات العشوائية المزدوجة
المحاضرة السابعة	التوزيع الأسي	المحاضرة الرابعة عشر	دوال المتغيرات العشوائية المزدوجة

المحاضرة الأولى

توزيع برنولي وتوزيع ذي الحدين (الثنائي)

المحاضرة الأولى: توزيع برنولي وتوزيع ذي الحدين (الثنائي)

تتناول هذه المحاضرة بنوع من التفصيل كل من التوزيعين الاحتماليين المتقطعين التاليين:

–توزيع برنولي "*Bernoulli distribution*": ويطلق عليه أيضا بـ "تجربة برنولي" او بـ "محاولة برنولي"، وتم تسميته بهذا الاسم نسبة لعالم الرياضيات السويسري "جاكوب برنولي" "*Jakob Bernoulli*" (1654-1705) الذي درسه، ودرس خصائصه مع نهاية القرن السابع عشر، الذي يصفه بتجربة لها نتيجتين فقط، النجاح "*Success*"، والفشل "*Failure*"؛

–توزيع ذي الحدين (الثنائي) "*Binomial distribution*": هو توزيع ساهم بجزء كبير في صياغته عالم الرياضيات الإنجليزي "إسحاق نيوتن" "*Isaac Newton*" (1643-1727)، ويقوم هذا التوزيع أيضا على عدد ثابت "n" من تجارب برنولي المستقلة.

المتغير العشوائي المتقطع (المنفصل) "*Discrete random variable*":

ان "X" متغير عشوائي متقطع إذا كان فضاء العينة، مجموعة قابلة للعد، سواء كانت مجموعة منتهية ام غير منتهية، او بمعنى آ خر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة متقطع تسمى متغيرا عشوائيا متقطعا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، دار ابن الاثير للطباعة والنشر في جامعة الموصل، 1990، ص41.

1-توزيع برنولي:

إذا كان لدينا تجربة عشوائية متعلقة بها حادثة معينة "A"، وعند قيامنا بهذه التجربة فان الحادثة "A" اما ان تحدث او لا تحدث، أي تحدث الحادثة "a"، لذا نلاحظ ان التجربة يكون لها دائما احدى النتيجتين الاتيتين:[1]

-اما ان تحدث الحادثة "A"، وتسمى هذه الحالة "حالة النجاح"؛

-اما لا تحدث الحادثة "A" أي تحدث الحادثة "a"، وتسمى هذه الحالة "حالة الفشل".

أي ان نجاح التجربة او فشلها يحدد بحدوث "A" او "a" على التوالي. وتسمى كل محاولة من محاولات إعادة مثل هذه التجربة بـ "محاولات برنولي" "*Bernoulli Trials*".[2]

1-1-التعريف:

تعريف توزيع برنولي:

نفترض ان المتغير "X" متغير عشوائي متقطع (منفصل)، ليكون "$X=1$" عندما تحدث "A" حالة النجاح، و"$X=0$" عندما تحدث "a" حالة الفشل، إذا كانت:

$-P(X=1)=p$;

$-P(X=0)=1-p=q$.

مما سبق نستطيع ان نكتب الصيغة الرياضية لتوزيع برنولي كما يلي:

$$P(X=x)=p^x q^{1-x}=p^x(1-p)^{1-x};x=0,1$$

ونقول عندئذ ان المتغير "X" يتبع توزيع برنولي بالمعلمة "p"، ونجد ان:

$$P(X=0)+P(X=1)=1$$

[1]-باسل يونس ذنون، "**الاحتمالية والمتغيرات العشوائية**"، دار الكتب للطباعة والنشر، الموصل، 1991، ص365.

[2]-نفس المرجع السابق، ص365.

نستطيع ان نلخص ما سبق في الجدول الموالي:

X	الاحتمال
0	q
1	p
$\sum$	1

والصيغة الرياضية لتوزيع برنولي تكتب كما يلي:

$$F(x)=\begin{cases}0;x<0\\q;0\leq x<1\\1;x\geq 1\end{cases}$$

1-2-المميزات العددية للتوزيع:

ان المميزات العددية لتوزيع برنولي تتمثل في:

-التوقع الرياضي:

$$E(X)=\sum x\times P(X=x)=0\times P(X=0)+1\times P(X=x)=0\times q+1\times p$$

$$E(X)=q$$

-التباين:

$$V(X)=E(X^2)-[E(X)]=\sum x^2P(X=x)-[p]^2$$

$$V(X)=0^2\times q=1^2\times p-p^2=p-p^2=p(1-p)$$

$$V(x)=p\times q$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{V(x)}=\sqrt{p\times q}$$

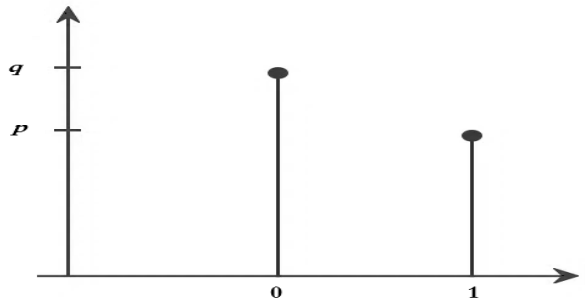
1-3-التمثيل البياني للتوزيع:

نقوم بالتمثيل البياني لدالة التوزيع ودالة الكثافة لتوزيع برنولي كما يلي:

-دالة توزيع برنولي الاحتمالية "*Probability Mass Function PMF*": يتم تمثيل دالة توزيع برنولي الاحتمالية "*PMF*" عند القيمتين 0، و1، كما يلي:

-عند القيمة "*X=1*": والذي يمثل حالة النجاح، تكون قيمة الاحتمال مساوية لـ "*p*"؛

-عند القيمة "*X=0*": والذي يمثل حالة النجاح، تكون قيمة الاحتمال مساوية لـ "*q*".



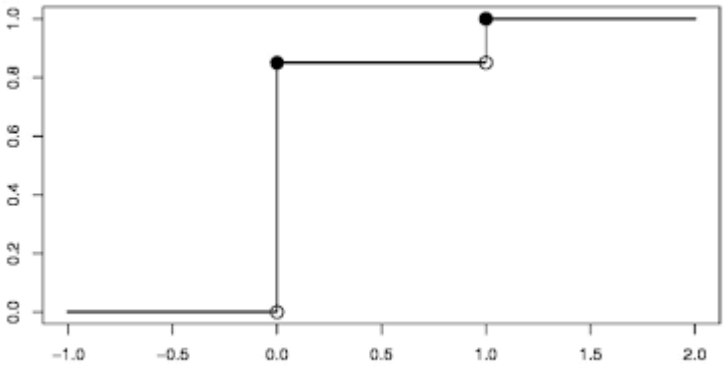
-دالة توزيع برنولي التراكمية "*Cumulative Distribution Function CDF*": يتم تمثيل دالة توزيع برنولي التراكمية "*CDF*" على شكل درج متزايدة (تصاعدية) عند القيم التالية كما يلي:

-القيمة صفر: عند الشرط $x<1$؛

-القيمة "*q*": عند الشرط $0\leq x<1$؛

-القيمة "*p*": عند الشرط $x\geq 1$.

فان التمثيل البياني لـ "*CDF*" كما يلي:



2-توزيع ذي الحدين (الثنائي):

يعتبر توزيع ذي الحدين أحد التوزيعات المتقطعة ذات أهمية تطبيقية كبيرة في الحياة العملية وخصوصا في موضوع الرقابة على جودة الإنتاج، وموضوع اختبارات النسب والنسب المئوية، ويمكن اعتبار هذا التوزيع حالة أكثر عمومية لتوزيع برنولي عندما يكون عدد المحالات أكثر من محاولة واحدة ($n \geq 1$).[1]

2-1-التعريف:

تعريف توزيع ذي الحدين (الثنائي):

إذا قمنا بتكرار تجربة برنولي "n" مرة، فسوف ننتقل الى قانون التوزيع ذي الحدين (الثنائي)، ويعرف القانون الاحتمالي لهذا التوزيع بالصيغة التالية:

$$P(X=x)=C_n^x \times p^x \times q^{n-x}; x=0,1,2,\ldots,n$$

ونكتب:

$$X \rightsquigarrow B(n,p)$$

أي ان "X" يتبع توزيع ذي الحدين (الثنائي) بالمعلمتين "n, p" علما ان:

$$C_n^x = \frac{n!}{x!(n-x)!}$$

2-2-المميزات العددية للتوزيع:

ان مميزات توزيع ذي الحدين (الثنائي) العددية هي:

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص233.

-التوقع الرياضي:

$$E(X)=n\times p$$

-التباين:

$$V(X)=n\times p(1-p)=n\times p\times q$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{V(x)}=\sqrt{n\times p\times q}$$

2-3-التمثيل البياني للتوزيع:

التمثيل البياني لدالة توزيع ذي الحدين (الثنائي)، يكون بتمثيل "X" على محور السينات الذي يمثل عدد النجاحات، وعلى محور العينات "$P(X)$" الذي يمثل الاحتمال الذي يقابل حدوث كل عدد من النجاحات.

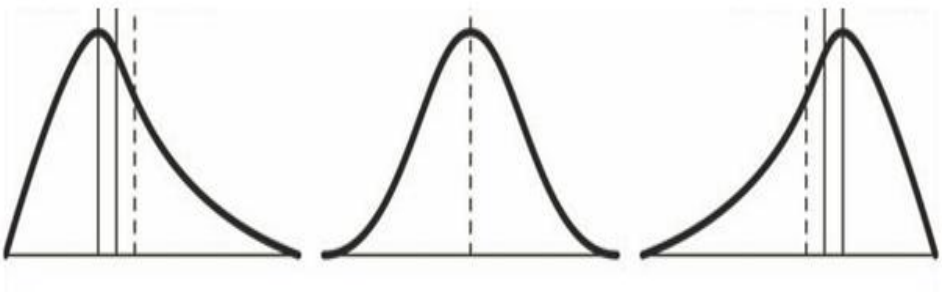
وبالتالي التمثيل البياني لدالة توزيع ذي الحدين (الثنائي) يتأثر بكل من عدد المحاولات "n" واحتمال النجاح "p" كما يلي:

-التناظر "*Symmetry*":

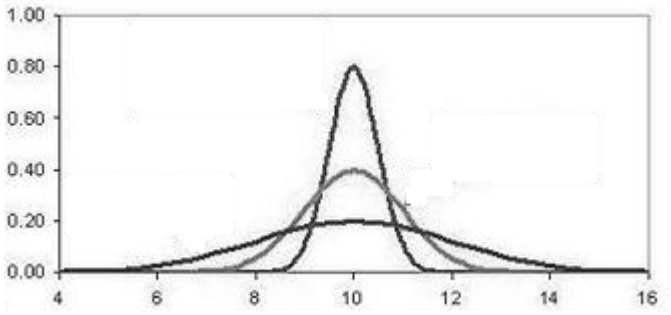
-"$p=0.5$": يكون التمثيل البياني لتوزيع ذي الحدين (الثنائي) متناظرا بالنسبة للوسط ويشبه رسم الجرس؛

-"$p<0.5$": يكون التمثيل البياني لتوزيع ذي الحدين (الثنائي) يميل نحو اليمين وتتركز الاحتمالات الكبيرة عند القيم الصغيرة لـ "X"؛

-"$p>0.5$": يكون التمثيل البياني لتوزيع ذي الحدين (الثنائي) يميل نحو اليسار وتتركز الاحتمالات الكبيرة عند القيم الكبيرة لـ "X"؛



وكلاحظة مهمة ان عدد المحاولات "n" كلما زاد عددها، فان الشكل البياني لتوزيع ذي الحدين (الثنائي) يقترب من الشكل البياني للتوزيع الطبيعي، ولو لم تكن قيمة الاحتمال "p" تساوي 0.5.



3-تمارين محلولة:[1]

التمرين الأول

ما هي التجارب التي نستطيع ان نقول عنها ان توزيع احتمالها يتبع توزيع برنولي؟

ان التجارب التي نستطيع القول عنها ان توزيع احتمالها يتبع توزيع برنولي، هي التجارب التي تتوفر فيها الشروط التالية:

-القيام بالتجربة مرة واحدة يسفر بنتيجتين اما النجاح او الفشل؛

-مجموع احتمال النتيجتين السابقتين تساوي الواحد الصحيح.

ومن بين التجارب المشهورة التي تتبع توزيع برنولي هي:

-ظهور الصورة في تجربة رمي قطعة نقود؛

-ظهور عدد يقبل القسمة على ثلاثة في تجربة رمي قطعة زهر النرد؛

-التنبؤ بسقوط الامطار يوم الغد؛

-نجاح أي مترشح في اجتياز أي مسابقة

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التمرين الثاني

في احدى المسابقات العلمية تم تقديم عشرة أسئلة للمترشحين، بحيث انه:

يحتوي كل سؤال على أربع خيارات منها خيار واحد صحيح، تم اختيار أحد المترشحين عشوائيا والذي أجاب على الأسئلة العشرة، فاذا اعتبرنا "X" متغير عشوائي يمثل عدد الإجابات الصحيحة.

المطلوب: ما هو احتمال ان يجيب المترشح على ثمانية أسئلة على الأقل بشكل صحيح من بين عشرة أسئلة؟

-حساب احتمال الإجابة على ثمانية أسئلة على الأقل بشكل صحيح من بين العشرة أسئلة:

ان احتمال الإجابة على أي سؤال بشكل صحيح هو:

$$p=\frac{1}{4}=0.25$$

ان احتمال الإجابة على أي سؤال بشكل خاطئ هو:

$$q=\frac{3}{4}=0.75$$

لدينا عشرة اسئلة:

$$n=10$$

$$X\rightsquigarrow B(10,\frac{1}{4})$$

$$P(X=x)=C_n^x(\frac{1}{2})^x(\frac{1}{2})^{n-x}\ ;\ x=0,1,2,\dots,n$$

$$P(X=x)=C_{10}^x(\frac{1}{2})^x(\frac{1}{2})^{10-x}\ ;\ x=0,1,2,\dots,10$$

$$P(X\geq 8)=P(X=8)+P(X=9)+P(X=10)$$

$$P(X=8)=C_{10}^8(\frac{1}{4})^8(\frac{3}{4})^{10-8}=\frac{10!}{8!2!}(\frac{1}{4})^8(\frac{3}{4})^2=0.00038$$

$$P(X=9)=C_{10}^9(\frac{1}{4})^9(\frac{3}{4})^{10-9}=\frac{10!}{9!1!}(\frac{1}{4})^9(\frac{3}{4})^1=0.00002$$

$$P(X=10)=C_{10}^{10}(\frac{1}{4})^{10}(\frac{3}{4})^{10-10}=\frac{10!}{10!0!}(\frac{1}{4})^{10}(\frac{3}{4})^0=0.0000009$$

$$P(X\geq 8)=0.00038+0.00002+0.0000009=0.0004009=0.04009\%$$

من خلال هذه النتيجة، نلاحظ ان احتمال ان يجيب المترشح عن ثمان أسئلة بشكل صحيح هو ضعيف جدا، ويعود هذا الى كون الاختيارات الممنوحة له في السؤال هي أربعة، لكن لو تم تخفيض عدد الخيارات فان احتمال الإجابة بشكل صحيح سيرتفع.

التمرين الثالث

اجتاز امتحان الإحصاء 3 سبعة طلبة، إذا افترضن ان احتمال النجاح في الامتحان هو 60%، ما هو احتمال:

أ-ان ينجح أربعة طلبة؛

ب-ان ينجح كل الطلبة.

أ-حساب احتمال ان ينجح أربعة طلبة في اجتياز امتحان الإحصاء 3:

ان هذه التجربة تتبع توزيع ذي الحدين (الثنائي) بالخصائص التالية:

$n=7$

$p=0.6$

$q=1-0.6=0.4$

إذا فان دالة توزيع ذي الحدين (الثنائي) بالمعطيات السابقة هي:

$X \rightsquigarrow B(n,p)$

$X \rightsquigarrow B(7,0.6)$

$P(X=x)=C_n^x \times p^x \times q^{n-x}; x=0,1,2,\dots,n$

$P(X=x)=C_7^x \times (0.6)^x \times (0.4)^{7-x}; x=0,1,2,\dots,7$

أي ان:

$P(X=4)=C_7^4 \times (0.6)^4 \times (0.4)^{7-4}$

$$P(X=4)=\frac{7!}{4!\,(7-4)!} \times (0.6)^4 \times (0.4)^3$$

$P(X=4)=0.2903$

ان احتمال نجاح أربعة طلبة من بين سبعة طلبة اجتازوا امتحان الإحصاء 3 هو 0.2903.

ب–حساب احتمال ان ينجح كل الطلبة في اجتياز امتحان الإحصاء 3:

$$P(X=7)=C_7^7\times(0.6)^7\times(0.4)^{7-7}$$

$$P(X=7)=\frac{7!}{7!\,(7-7)!}\times(0.6)^7\times(0.4)^0$$

$$P(X=7)=0.028$$

ان احتمال نجاح كل الطلبة الذين اجتازوا امتحان الإحصاء 3 هو 0.028.

التمرين الرابع

قمنا بتجربة القاء قطعة نقود متزنة ست مرات، قم بحساب:

أ–احتمال الحصول على صورتين؛

ب–متوسط عدد مرات ظهور الصورة خلال ست مرات.

أ–حساب احتمال الحصول على صورتين:

لدينا:

$$n=6$$

$$p=q=\frac{1}{2}$$

اذن:

$$X\rightsquigarrow B(6,\frac{1}{2})$$

$$P(X=x)=C_n^x(\frac{1}{2})^x(\frac{1}{2})^{n-x}\;;\;x=0,1,2,\ldots,n$$

$$P(X=x)=C_6^x(\frac{1}{2})^x(\frac{1}{2})^{6-x}\;;\;x=0,1,2,\ldots,6$$

$$P(X=2)=C_6^2(\frac{1}{2})^2(\frac{1}{2})^{6-2}=\frac{6!}{2!4!}(\frac{1}{2})^2(\frac{1}{2})^4=15(0.25)(0.0625)$$

$$P(X=2)=0.2324$$

ان احتمال الحصول على صورتين خلال تجربة رمي قطعة نقو متزنة ست مرات هو 0.2324.

ب–حساب متوسط عدد مرات ظهور الصورة خلال ست مرات:

$$E(X)=n\times p$$

$$E(X)=6\times(\frac{1}{2})=3$$

المحاضرة الثانية

توزيع بواسون

المحاضرة الثانية: توزيع بواسون

ان:

–توزيع بواسون "*Poisson distribution*": هو توزيع قام بتطويره عالم الرياضيات والفيزياء الفرنسي "سيميون دينيس بواسون" "*Siméon Denis Poisson*" (1781–1840)، مع أوائل القرن التاسع عشر، من اجل وصف عدد مرات فوز المقامر في لعبة قمار على المدى الطويل، ويستخدم هذا التوزيع لحساب احتمال وقوع عدد معين من الأحداث في فترة زمنية محددة، مع توفر الاستقلالية بين الاحداث وبمعدل ثابت.

المتغير العشوائي المتقطع (المنفصل) "*Discrete random variable*":

ان "X" متغير عشوائي متقطع إذا كان فضاء العينة، مجموعة قابلة للعد، سواء كانت مجموعة منتهية ام غير منتهية، او بمعنى آ خر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة متقطع تسمى متغيرا عشوائيا متقطعا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-توزيع بواسون:

في الحياة العملية أحيانا ما نقابل بعض الظواهر التي ينطبق عليها شروط توزيع ذي الحدين (الثنائي)، ولكن الحوادث تكون نادرة الوقوع وهذا يعني ان احتمال النجاح يكون صغيرا جدا ويقترب من الصفر، وعليه فانه يمكن القول ان "$np=\lambda$"، "λ" هو مقدار ثابت، وبذلك يكون احتمال الفشل كبير أي انه يقترب من الواحد،[1] وبذلك تكون شروط هذا التوزيع كالاتي:[2]

-ان يكون احتمال النجاح "p" ثابت وكذلك احتمال الفشل "q" في كل محاولة؛

-ان يكون احتمال النجاح صغيرا ويقترب من الصفر، واحتمال الفشل يقترب من الواحد الصحيح؛

-ان يكون عدد المحاولات كبيرا جدا حيث ان "$np=\lambda$" مقدار ثابت.

ولتوزيع بواسون تطبيقات واسعة فهو يقدم بشكل عام نموذجا للمعلومات الإحصائية التي تأخذ شكل تعداد لحوادث نادرة الوقوع، حيث يمثل المتغير العشوائي "X" عدد الحوادث النادرة الملحوظة في وحدة قياس معينة، زمنا كانت ام مساحة ام حجما بينما يمثل "λ" معدل او متوسط عدد مرات ظهور تلك الاحداث في وحدة القياس. وتمثل الأمثلة التالية الحالات التي تطبق فيها توزيع بواسون:[3]

-عدد المكالمات الهاتفية التي ترد الى مقسم خلال فترة زمنية معينة؛

-عدد الجسيمات الصادرة في الثانية عن مادة مشعة؛

-عدد الأخطاء المطبعية في صفحة من صفحات كتاب معين؛

-عدد حوادث المرور في مدينة كبيرة خلال يوم معين؛

-عدد البكتيريا الموجودة في ميليمتر مكعب واحد من وعاء يحتوي على سائل معين.

---

[1]-رامز قدسية، "**الاحتمالات والاحصاء**"، منشورات الجامعة الافتراضية السورية، 2018، ص147.

[2]-نفس المرجع السابق، ص147.

[3]-نفس المرجع السابق، ص147.

1-التعريف:[1]

تعريف توزيع بواسون:

يستخدم هذا التوزيع لتحديد احتمال عدد معين من الاحداث (النجاحات) في وحدة من الزمن، وذلك عندما تكون الاحداث مستقلة عن بعضها البعض ويبقى متوسطها ثابتا لوحدة من الزمن، مثل عدد حوادث السيارات خلال السنة، عدد المكالمات خلال الشهر، ... .

شروط هذا التوزيع هي:

-ان يكون احتمال النجاح ثابتوكذلك احتمال الفشل في كل محاولة لهما بالرمز "q"، "p"؛

-ان يكون احتمال النجاح صغيرا ويقترب من الصفر واحتمال الفشل يقترب من الواحد صحيح؛

-ان يكون عدد المحاولات كبير جدا على ان تكون:

$$n \times p = \lambda$$

مقدار ثابت.

يمكن صياغة قانون توزيع بواسون كما يلي:

$$p(X) = \frac{\lambda^x e^{-x}}{x!} ; x = 0,1,2, \ldots, n.$$

حيث ان:

-X : العدد المعين من النجاحات؛

-$p(X)$: احتمال عدد "x " من النجاحات؛

-λ : متوسط عدد النجاحات في وحدة الزمن؛

-e : 2.71828.

[1]-بوزنورة اسماء، "**الإحصاء (02) –دروس وتمارين–**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 03، 2024/2023، ص74-76.

نلاحظ ان هذا التوزيع معرف فقط في المعلمة "λ " أي يكفي معرفتها لتطبيق هذا التوزيع ويعرف كما يلي:

$X \rightsquigarrow P(\lambda)$

2-المميزات العددية للتوزيع:

ان المميزات العددية لتوزيع بواسون تتمثل في:

-التوقع الرياضي:

التوقع الرياضي للمتغير العشوائي الخاضع لتوزيع بواسون هو معلمة هذا التوزيع أي:

$E(X)=\lambda$

-التباين:

$V(X)=\lambda$

-الانحراف المعياري:

$\sigma_X=\sqrt{V(x)}=\sqrt{\lambda}$

3-التمثيل البياني للتوزيع:

ان توزيع بواسون من بين التوزيعات المتقطعة او المنفصلة المشهورة، ولكي نستطيع ان نمثله بيانيا نقوم بما يلي:

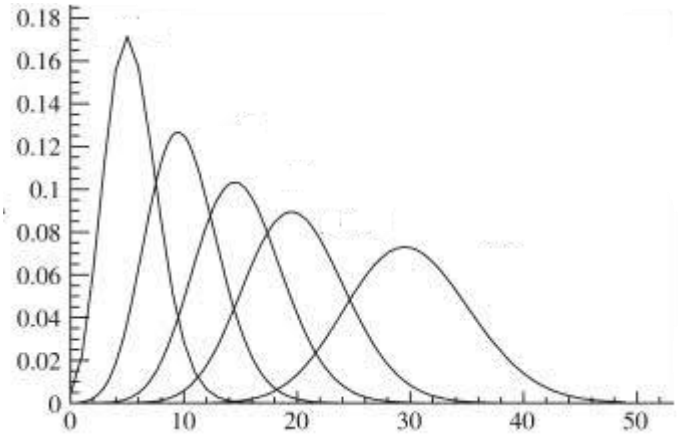
-على المحور الافقي (السيني): يتم تحديد عدد مرات وقوع الحدث، ويأخذ قيما صحيحة فقط؛

-على المحور العمودي (العيني): يتم تحديد قيم الاحتمالات لكل قيمة من قيم المحور الافقي.

ان المعامل "λ " يستطيع ان يؤثر على التمثيل البياني لتوزيع بواسون كما يلي:

-عندما تكون قيمة "λ " صغيرة: يكون التمثيل البياني لتوزيع بواسون ملتويا جهة اليمين "*Positively Skewed*"، فتتركز الاحتمالات الكبيرة بالقرب من القيم الصغيرة؛

-وكلما ازدادت قيمة "λ ": كلما أصبح التمثيل البياني لتوزيع بواسون أكثر تناظرا ويبدأ يقترب من شكل منحنى الجرس "*Bell Shape*"، وبالتالي يقترب من الشكل البياني للتوزيع الطبيعي.



4-تمارين محلولة:[1]

التمرين الأول

إذا كان عدد حوادث السيارات في مدينة معينة يتبع توزيع بواسون بمتوسط خمس حوادث خلال أسبوع، إذا افترضنا ان "X" متغير عشوائي يمثل حوادث السيارات خلال أسبوع.

المطلوب:

أ-حدد شكل دالة الاحتمال $F(X)$ لهذا المتغير؛

ب-ما هو احتمال حدوث حادث واحد على الأقل خلال أسبوع؟

ج-ما هو احتمال حدوث ثلاث حوادث على الأكثر خلال أسبوع؟

د-احسب التوقع الرياضي، التباين والانحراف المعياري لعدد الحوادث؛

هـ-حدد شكل التوزيع.

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

أ–تحديد شكل دالة الاحتمال *F(X)*:

بما ان متوسط عدد حوادث السيارات خلال أسبوع هو $x=5$ وبالتالي تكون دالة الاحتمال:

$$F(X)=\frac{\lambda^x e^{-x}}{x!};x=0,1,2,\ ...,n$$

$$F(X)=\frac{\lambda^5 e^{-5}}{5!};x=0,1,2,\ ...,5$$

ب–حساب احتمال حدوث حادث واحد على الأقل خلال أسبوع:

$$F(x\geq 1)=1-F(X)$$

$$F(x\geq 1)=1-\frac{5^0 e^{-5}}{0!}$$

$$F(x\geq 1)=0.99$$

ج–حساب احتمال حدوث ثلاث حوادث على الأكثر خلال أسبوع:

$$F(x\leq 3)=F(3)+F(2)+F(1)+F(0)$$

$$F(x\leq 3)=\frac{5^3 e^{-5}}{3!}+\frac{5^2 e^{-5}}{2!}+\frac{5^1 e^{-5}}{1!}+\frac{5^0 e^{-5}}{0!}$$

$$F(x\leq 3)=0.25$$

د–حساب التوقع الرياضي، التباين والانحراف المعياري لعدد الحوادث:

–التوقع الرياضي:

$$E(X)=\lambda=5$$

–التباين:

$$V(X)=\lambda=5$$

–الانحراف المعياري:

$$\sigma_X=\sqrt{V(X)}=\sqrt{\lambda}=\sqrt{5}=2.23$$

هـ-تحديد شكل التوزيع:

ان توزيع بواسون دائما موجب الالتواء.

التمرين الثاني

اذا كان متوسط عدد مستخدمي الصراف الالي في احد البنوك هو خمسة اشخاص كل نصف ساعة.

أحسب:

-الاحتمالات التالية لأعداد الواصلين كل نصف ساعة بان يكون:

أ-عشرة اشخاص؛

ب-يقل عن ثلاثة اشخاص؛

ج-يتراوح العدد بين أربعة اشخاص وثمانية اشخاص.

أ-حساب احتمالات عدد الواصلين كل نصف ساعة بان يكون عشرة اشخاص:

معدل الوصول كل نصف ساعة هو:

$\lambda=5$

اذن:

$$F(X)=\frac{\lambda^x e^{-x}}{x!};x=0,1,2,\ ...,n$$

$$F(X)=\frac{\lambda^{10} e^{-5}}{10!};x=0,1,2,\ ...,10$$

$$F(X=10)=\frac{5^{10} e^{-5}}{10!}$$

$$F(X=10)=0.01813$$

ان احتمال عدد الواصلين كل نصف ساعة يكون عشرة اشخاص هو 0.01813.

ب–حساب احتمالات عدد الواصلين كل نصف ساعة بان يكون اقل من ثلاث اشخاص:

$$F(X<3)=P(X=0)+P(X=1)+P(X=2)$$

$$F(X<3)=\frac{5^0e^{-5}}{0!}+\frac{5^1e^{-5}}{1!}+\frac{5^2e^{-5}}{2!}$$

$$F(X<3)=0.12465$$

ان احتمال عدد الواصلين كل نصف ساعة يكون اقل من ثلاثة اشخاص هو 0.12465.

ج–حساب احتمالات عدد الواصلين كل نصف ساعة بان يكون يتراوح بين أربعة وثمانية اشخاص:

$$F(4\leq X\leq 8)=P(X=4)+P(X=5)+P(X=6)+P(X=7)+P(X=8)$$

$$F(4\leq X\leq 8)=0.66688$$

ان احتمال عدد الواصلين كل نصف ساعة يكون يتراوح بين أربعة اشخاص وثمانية اشخاص هو 0.66688.

التمرين الثالث

لوحظ ان هناك بالمعدل 2 بكتيريا لكل سنتيمتر مكعب من الماء المأخوذ من مستنقع معين، في نموذج من ماء ذلك المستنقع حجمه 2 سنتيمتر مكعب، احسب احتمال ان يحتوي هذا النموذج:

أ–في الأكثر خمس بكتيريات؛

ب–في الأقل أربع بكتيريات.

أ–حساب احتمال ان يحتوي هذا النموذج في الأكثر خمس بكتيريات:

ان "X" يمثل عدد البكتيريات الموجودة في النموذج، ولما كان "X" يأخذ قيما غير محدودة من المعدودات في وحدة حجمية، فمن المعقول ان نفترض ان "X" يتبع توزيع بواسون، ولما كان هناك بالمعدل 2 بكتيريا لكل سنتيمتر مكعب من الماء، فيكون هناك بالمعدل أربع بكتيريات لكل 2 سنتيمتر مكعب من الماء، لهذا نختار 4.

$$F(X)=\frac{\lambda^x e^{-x}}{x!};x=0,1,2,\ ...,n$$

$$F(X)=\frac{4^x e^{-4}}{x!};x=0,1,2,\ ...$$

$F(X\leq 5)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)$

$F(X\leq 5)=0.018+0.073+0.147+0.195+0.195+0.156$

$F(X\leq 5)=0.784$

احتمال ان يحتوي هذا النموذج في الأكثر خمس بكتيريات هو 0.784.

ب-حساب احتمال ان يحتوي هذا النموذج في الأقل أربع بكتيريات:

$F(X\geq 4)=1-P(X<4)$

$F(X\geq 4)=1-P(X=0)-P(X=1)-P(X=2)-P(X=3)$

$F(X\geq 4)=0.567$

احتمال ان يحتوي هذا النموذج في الأقل أربع بكتيريات هو 0.567.

التمرين الرابع

إذا كان عدد الأخطاء المطبعية في كتاب يتكون من 600 صفحة هو 50 خطا، فاذا كانت الأخطاء تتوزع توزيعا عشوائيا.

فما احتمال إذا اختيرت 10 صفحات عشوائيا الا تحتوي على أخطاء.

ان "X" يمثل عدد الأخطاء في كل صفحة وان عدد المحاولات (الصفحات) تمثل سلسلة من محاولات برنولي، عددها "$n=100$"، نسبة الخطا (النجاح) هي:

$p=\frac{50}{600}$

$p=0.083$

وعليه فان:

$\lambda=np$

$\lambda=10(0.083)$

$\lambda=0.83$

وبالتالي فان "X" يتبع التوزيع الطبيعي.

وان:

$$F(X)=\frac{\lambda^x e^{-x}}{x!};x=0,1,2,\ ...,n$$

$$F(X=0)=\frac{0.83^0 e^{-0.83}}{0!}$$

$$F(X=0)=0.436$$

ان احتمال إذا اختيرت 10 صفحات عشوائيا الا تحتوي على أخطاء هو 0.436.

المحاضرة الثالثة

التوزيع الهندسي

المحاضرة الثالثة: التوزيع الهندسي

ان:

-التوزيع الهندسي "*Geometric distribution*": تطور مع تطور نظرية الاحتمالات، ولا ينسب الى عالم معين، وفي بعض الأحيان ينسب الى عالم الرياضيات الأمريكي "ويندل فوري" "*Wendell F. Forrey*"، وتعود تسمية "الهندسي" الى ان دالة التوزيع الهندسي تتبع في صيغتها المتتالية الهندسية. وهو يمثل حالة خاصة من تجربة برنولي المتكررة، بتجارب لها نفس الاحتمال ومستقلة فيما بينها، اذ ان له علاقة أيضا بتوزيع ذي الحدين (الثنائي).

المتغير العشوائي المتقطع (المنفصل) "*Discrete random variable*":

ان "X" متغير عشوائي متقطع إذا كان فضاء العينة، مجموعة قابلة للعد، سواء كانت مجموعة منتهية ام غير منتهية، او بمعنى آ خر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة متقطع تسمى متغيرا عشوائيا متقطعا.[1]

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع الهندسي:

يعد هذا التوزيع من التوزيعات المهمة في التطبيقات الإحصائية، خاصة المتعلقة بدراسة الإحصاء السكاني، حيث يستخدم في دراسة معدلات النمو ومعدلات الوفيات والولادة، ...، فاذا كان هناك تجربة إحصائية تتألف من متوالية من محاولات برنولي المستقلة، وكانت نتيجة كل محاولة من هذه المحاولات اما نجاح او فشل، وستجرى التجربة لحين الحصول على اول حالة نجاح، وعلى فرض ان احتمال النجاح ثابت في كل محاولة وليكن "p" (الحظ هنا ان عدد المرات "n" متغير عشوائي بينما في حالة ذي الحدين (الثنائي)، عدد المرات ثابت)، وكان المتغير العشوائي "X" يمثل عدد المحاولات المطلوبة لوقوع اول حالة نجاح فان "X" يتوزع وفق التوزيع الهندسي، ومن الأمثلة على ذلك، القاء عملة نقدية معدنية تكرارا حتى ظهور اول صورة، القاء مكعب نرد حتى الرقم 4، سحب عناصر من صندوق به عناصر فاسدة وأخرى صالحة بشكل متتالي ومع الإعادة حتى الحصول على عنصر فاسد، ...[1]

اذن يستخدم التوزيع الهندسي من اجل معرفة عدد المحاولات اللازمة من اجل الحصول على اول نجاح من سلسلة تجارب تتوفر فيها شرط الاستقلالية فيما بينها، ونستطيع ان نذكر اهم استعمالاته فيما يلي:

- **-لمراقبة جودة المنتجات**، أي تحديد عدد عمليات الفحص الواجبة من اجل على اول وحدة منتجة من المنتوج تكون فيها عيب او فاسدة؛
- **-مجال الاتصالات**، لمعرفة عدد المحاولات لأجل ارسال البيانات بطريقة ناجحة عبر شبكة الاتصالات؛
- **-عالم الاحياء**، لوصف بعض التوزيعات التي تمثل الاطوال لبعض التراكيب البيولوجية؛
- **-المراهنة والألعاب**، لتحديد عدد مرات الرمي من اجل الحصول على نتيجة معينة؛
- **-البيع والتسويق للمنتجات**، لتحديد عدد المحاولات اللازمة لبيع وتسويق وحدة واحدة.

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، دار الحكمة، 2000، ص334.

1-التعريف:[1]

تعريف التوزيع الهندسي:

نقول ان المتغير العشوائي "X" يتبع التوزيع الهندسي إذا كان قانون احتماله يكتب وفق الشكل التالي:

$$P(X=x)=p(1-p)^{x-1};x=1,2,...,n$$

نطبق هذا القانون إذا كان لدينا تجارب نقوم بتكرارها من اجل الحصول على اول نجاح، فمثلا نقوم بفحص السلع لمنتج معين ويكون "X" عدد السلع المفحوصة حتى نصل الى اول سلعة تالفة، فاحتمال النجاح هو "p" واحتمال الفشل هو "$q=1-p$".

إذا كان "X" متغيرا عشوائيا يمثل عدد المحاولات المطلوبة للحصول على اول حالة نجاح، في متوالية من المحاولات المستقلة لتجربة ما، وكان احتمال النجاح هو "p"، وهو ثابت من محاولة لأخرى، فان "X" يتوزع وفق التوزيع الهندسي بدالة كتلة احتمال معرفة بالصيغة الاتية:[2]

$$P(X=x) = \begin{cases} pq^{x-1}, & x=1,2,..., 0<p<1 \\ 0 & or \end{cases}$$

ويرمز لذلك بالرمز:

$$X \rightsquigarrow Ge(p)$$

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع الهندسي تتمثل في:

-التوقع الرياضي:

$$E(X)=\frac{1}{p}$$

[1]-عبدلي ادريس، "**الإحصاء 03**"، دروس عبر الخط في مقياس، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة البليدة 02، 2024/2023، ص23-24.

[2]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص335.

-التباين:

$$V(X)=\frac{q}{p^2}=\frac{1-p}{p^2}$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{V(X)}=\sqrt{\frac{q}{p^2}}=\sqrt{\frac{1-p}{p^2}}=\frac{\sqrt{1-p}}{p}$$

3-التمثيل البياني للتوزيع:

يتميز التمثيل البياني للتوزيع الهندسي بالمميزات التالية:

-ان التمثيل البياني للتوزيع الهندسي يكون ملتويا نحو اليمين "*Skewed to the right*" على الدوام؛

-منواله الذي يمثل أكثر قيمة تكرارا او احتمالا هي كل دائما عبارة عن قيمة الواحد الصحيح (1)؛

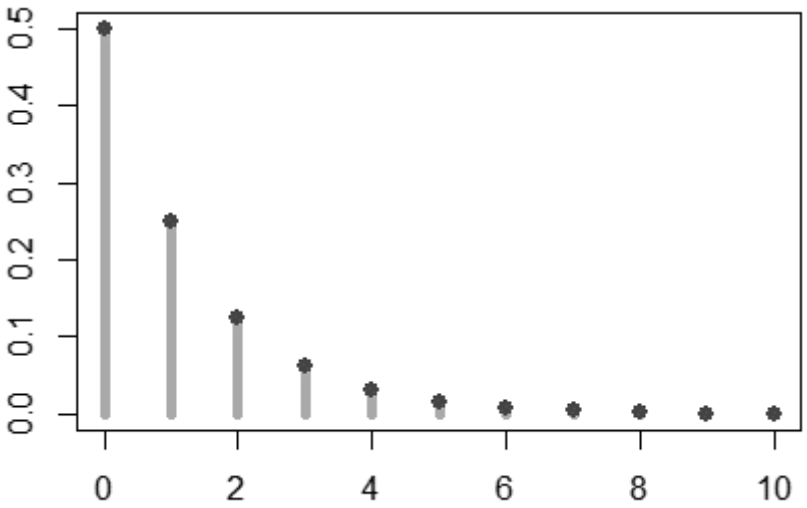
-الدالة التراكمية شكلها دائما على شكل درج يقترب من الواحد كلما زادت عدد المحاولات؛

-هذا التوزيع يستخدم لتوضيح انه كلما زاد عدد المحاولات الفاشلة، فانه يقل احتمال وقوع اول حالة نجاح.

لكي نستطيع ان نرسم التمثيل البياني للتوزيع الهندسي يجب:

-على المحور الافقي، نمثل عدد المحاولات حتى الوصول الى اول نجاح؛

-على المحور العمودي، نمثل قيم الاحتمالات المقابلة لعدد المحاولات.



4-تمارين محلولة:[1]

التمرين الأول

صياد يصيب الهدف باحتمال 0.8.

أ-ما هو احتمال ان يرمي على هدف معين خمس مرات حتى يصيب الهدف لأول مرة، ثم احسب التوقع الرياضي والتباين؟

ب-ما هو احتمال ان يصيب الهدف لأول مرة بعد ثلاث رميات على الأكثر؟

أ-احتمال ان يرمي على هدف معين خمس مرات حتى يصيب الهدف لأول مرة والتوقع الرياضي والتباين:

لنعتبر ان "X" متغير عشوائي يمثل عدد مرات الرمي نحو الهدف حتى اصابته لأول مرة، أي ان:

$$X \rightsquigarrow G(0.8)$$

وعليه القانون الاحتمالي يكتب كما يلي:

$$P(X=x)=p(1-p)^{x-1};x=1,2,...,n$$

$$P(X=5)=0.8(1-0.8)^{5-1};x=1,2,...,5$$

$$P(X=5)=0.8(0.2)^4$$

$$P(X=5)=0.00128$$

التوقع الرياضي:

$$E(X)=\frac{1}{p}$$

$$E(X)=\frac{1}{0.8}$$

$$E(X)=1.25$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التباين:

$$V(X)=\frac{q}{p^2}$$

$$V(X)=\frac{1-p}{p^2}$$

$$V(X)=\frac{1-0.8}{0.8^2}$$

$$V(X)=0.5$$

ب–حساب احتمال ان يصيب الهدف لأول مرة بعد ثلاث رميات على الأكثر:

$$P(X\leq 3)=P(X=3)+P(X=2)+P(X=1)$$

$$P(X=x)=p(1-p)^{x-1};x=1,2,\ldots,n$$

$$P(X=1)=0.8(1-0.8)^{1-1}=p=0.8$$

$$P(X=2)=0.8(1-0.8)^{2-1}=0.8(0.2)^{1}=0.8\times 0.2=0.16$$

$$P(X=3)=0.8(1-0.8)^{3-1}=0.8(0.2)^{2}=0.032$$

$$P(X\leq 3)=0.8+0.16+0.032$$

$$P(X\leq 3)=0.992$$

التمرين الثاني

نريد القاء مكعب زهرة نرد ونعيد العملية حتى يظهر الرقم واحد "1"، فاذا كان المتغير العشوائي "X" يمثل عدد المحاولات المطلوبة حتى ظهور الرقم واحد "1" لأول مرة، فأوجد:

أ–دالة الاحتمال للمتغير العشوائي "X"؛

ب–احتمال الحصول على الرقم واحد "1" في المحاولة الثالثة؛

ج–القيمة المتوقعة والتباين للمتغير العشوائي "X".

أ– دالة الاحتمال للمتغير العشوائي X:

$$P(X=x) = \begin{cases} (\frac{5}{6})(\frac{1}{6})^{x-1}, & x=1,2,\dots \\ 0 & or \end{cases}$$

ب–احتمال الحصول على الرقم واحد "1" في المحاولة الثالثة:

$$P(X=3)=(\frac{5}{6})(\frac{1}{6})^{3-1}$$

$$P(X=3)=\frac{25}{216}$$

ج–القيمة المتوقعة والتباين للمتغير العشوائي X:

التوقع:

$$E(X)=\frac{1}{p}$$

$$E(X)=\frac{1}{\frac{1}{6}}$$

$$E(X)=6$$

التباين:

$$V(X)=\frac{q}{p^2}$$

$$V(X)=\frac{\frac{5}{6}}{(\frac{1}{6})^2}$$

$$V(X)=30$$

التمرين الثالث

وجدت شركة للإلكترونيات ان 30% من المتقدمين للعمل بها يحملون شهادة في الحاسوب، وان المتقدمين يتم اختيارهم بطريقة عشوائية وبشكل متتالي.

أ-ما احتمال ان اول شخص تتم مقابلته ويحمل شهادة في الحاسوب هو الشخص الخامس؛

ب-إذا كانت كل مقابلة تكلف 15 دينار اوجد القيمة المتوقعة والتباين لإجمالي تكاليف المقابلة حتى يتم الحصول على العدد المطلوب من الأشخاص للوظائف الشاغرة.

أ-حساب احتمال ان اول شخص تتم مقابلته ويحمل شهادة في الحاسوب هو الشخص الخامس:

بفرض ان المتغير العشوائي "X" يرمز لعدد المحاولات المطلوب للحصول على اول شخص تتم مقابلته ويحمل شهادة في الحاسوب، وعليه فان المتغير العشوائي "X" يتوزع وفق التوزيع الهندسي بدالة احتمال لها الصيغة الاتية:

$$P(X=x) = \begin{cases} (0.3)(0.7)^{x-1}, & x=1,2,\dots \\ 0 & or \end{cases}$$

$$P(X=3)=(0.3)(0.7)^{5-1}$$

$$P(X=3)=0.072$$

ب-ايجاد القيمة المتوقعة والتباين لإجمالي تكاليف المقابلة:

بفرض ان "C" ترمز لإجمالي التكلفة وعليه فان:

$$C=15X$$

ومنها نجد ان:

$$E(C)=15E(X)$$

$$E(C)=15(\frac{1}{9})$$

$$E(C)=50$$

وان:

$$V(C)=(15)^2V(X)$$

$$V(C)=(15)^2(\frac{q}{p^2})$$

$$V(C)=(15)^2(\frac{0.7}{0.3^2})$$

$$V(C)=1750$$

التمرين الرابع

كيس يحتوي على 8 كرات بيضاء و4 كرات سوداء سحبت من الكيس كرة واحدة مع الإعادة والمطلوب:

أ-إذا كان "X" يمثل عدد مرات سحب كرة بيضاء، اوجد القيمة المتوقعة والتباين للمتغير العشوائي "X"؟

ب-احتمال ظهور كرة بيضاء في اول سحب من السحبة الخامسة؟

أ-ايجاد القيمة المتوقعة والتباين للمتغير العشوائي X:

نلاحظ ان احتمال سحب كرة بيضاء هو:

$$P(x)=\frac{8}{12}$$

$$P(x)=\frac{2}{3}$$

وعليه وبما ان "X" له توزيع هندسي فان:

$$P(x)=\begin{cases}(\frac{2}{3})(1-\frac{2}{3})^{x-1}, & x=1,2,\dots \\ 0 & \text{or}\end{cases}$$

وتكون القيمة المتوقعة للمتغير "X" هو:

$$E(X)=\frac{1}{p}$$

$$E(X)=\frac{1}{\frac{2}{3}}$$

$$E(X)=\frac{3}{2}$$

اما التباين فيكون:

$$V(X)=\frac{1-p}{p^2}$$

$$V(X)=\frac{1-\frac{2}{3}}{(\frac{2}{3})^2}$$

$$V(X)=\frac{3}{4}$$

ب-احتمال ظهور كرة بيضاء في اول سحب من السحبة الخامسة:

$$P\left(5,\frac{2}{3}\right)=\frac{2}{3}(1-\frac{2}{3})^{5-1}$$

$$P\left(5,\frac{2}{3}\right)=\frac{2}{3}(\frac{1}{3})^{4}$$

$$P\left(5,\frac{2}{3}\right)=\frac{2}{243}$$

المحاضرة الرابعة

التوزيع فوق الهندسي

المحاضرة الرابعة: التوزيع فوق الهندسي

ان:

-التوزيع فوق الهندسي "*Hypergeometric distribution*": هو توزيع لم يرتبط تطوره بباحث او عالم معين، ولا بفترة محددة، ويعتبر تطور وظهور هذا التوزيع كضرورة لمعالجة بعض الحالات الخاصة لتوزيع ذي الحدين (الثنائي)، خصوصا في حالة عدم تحقق شرط استقلالية الاحداث فيما بينها، وتظهر هذه الحالة عند السحب بدون ارجاع لأفراد عينة الدراسة. وهو يرتبط بكل من التوزيع ذي الحدين (الثنائي)، وتوزيع برنولي، فهو توزيع احتمالي يجسد سحب عدد معين من السحوبات الناجحة (او غير الناجحة، الفاشلة) من مجتمع (مجموعة) محدود (ة).

المتغير العشوائي المتقطع (المنفصل) "*Discrete random variable*":

ان "X" متغير عشوائي متقطع إذا كان فضاء العينة، مجموعة قابلة للعد، سواء كانت مجموعة منتهية ام غير منتهية، او بمعنى آ خر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة متقطع تسمى متغيرا عشوائيا متقطعا.[1]

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع فوق الهندسي:

لنفترض وجود مجتمع محدود يتضمن نوعين من العناصر فقط، وان عينة بحجم ثابت قد تم اختيارها بشكل متتال، وبدون إعادة، وان الهدف من هذه التجربة هو معرفة عدد عناصر احد النوعين بالعينة، فمثلا، اذا كان لدينا صندوق يحتوي على أربعة كرات بيضاء وخمس كرات أخرى حمراء، وانه تم سحب ثلاثة كرات وبدون إعادة من ذلك الصندوق، وكان المتغير العشوائي "X" يمثل عدد الكرات الحمراء المسحوبة، فان المتغير العشوائي "X" يتوزع وفق التوزيع فوق الهندسي، او مثلا، من الممكن وجود مجموعة من الأشخاص تتكون من عشرون رجلا، وعشرة نساء، اختيرت عينة منها حجمها خمسة عشر شخصا وبدون إعادة، وكان المتغير العشوائي "X" يمثل عدد النساء بالعينة المختارة، فان "X" يتوزع وفق التوزيع فوق الهندسي، ولوضع مثل هذه النوعية من التجارب في صورة دالة احتمالية لنفترض ان صندوق يحتوي على "N" نضيدة جافة، منها "M" نضيدة معيبة، و"N-M" نضيدة سليمة.[1] فاذا تم اختيار عينة عشوائية حجمها "k" من ذلك الصندوق وبدون إعادة، وكان المتغير العشوائي "X" يمثل عدد النضائد المعيبة بالعينة التي تم اختيارها فان هناك:[2]

- $\binom{N}{k}$ من الطرائق التي يمكن بها اختيار "k" نضيدة من بين "N" نضيدة؛
- $\binom{N}{x}$ من الطرائق التي يمكن بها اختيار "x" نضيدة معيبة من بين "M" نضيدة معيبة؛
- $\binom{N-M}{k-x}$ من الطرائق التي يمكن بها اختيار "$k-x$" نضيدة سليمة من بين "N-M" نضيدة سليمة.

ان التوزيع فوق الهندسي يستخدم في اغلب الأحيان في:

-مراقبة مدى مطابقة وجودة كمية من المنتجات، بواسطة اخذ عينة من شحنة من المنتجات لفحص القطع التالفة بدون ان يتم اعادتها ضمن الشحنة؛

-العاب اليانصيب، لمعرفة احتمال الفوز حيث في هذا النوع من الألعاب الأرقام لا تعاد من اجل السحب؛

-تحديد فصيلة معينة من الحيوانات في منطقة معينة من اجل دراسة في عالم الاحياء.

---

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص328.

[2]-نفس المرجع السابق، ص328.

1-التعريف:[1]

تعريف التوزيع الهندسي:

نقول ان المتغير العشوائي "X" يتبع التوزيع فوق الهندسي بمعالم "N, k, n"، إذا كان قانونه الاحتمالي يكتب كما يلي:

$$P(X=x)=\frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$$

يستعمل هذا التوزيع إذا كان لدينا صندوق يحتوي على "N" سلعة، حيث يوجد من بينها "k" سلعة تالفة، ونقوم بسحب عينة حجمها "n" دفعة واحدة، فاذا اعتبرنا "X" متغير عشوائي يمثل عدد السلع التالفة المسحوبة، في هذه الحالة فان "X" يتبع التوزيع فوق الهندسي، ونكتب:

$X \rightsquigarrow H(N,k,n)$

اما إذا كان السحب بالإرجاع، يعني نريد سحب ثلاث منتجات، لكن طريقة السحب تكون من خلال سحب المنتوج الأول، ثم نرجعه، ثم نسحب مرة ثانية، ثم نرجعه، ثم نسحب مرة ثالثة ثم نرجعه، في هذه الحالة نستعمل التوزيع الثنائي.

من كل ما سبق نستطيع ان نكتب دالة الاحتمال كما يلي:

$$\mathrm{P(X=x)}=\begin{cases}\dfrac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}}, & x=0,1,2,\ldots,k \\ 0 & or\end{cases}$$

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع فوق الهندسي تتمثل في:

-التوقع الرياضي:

$$E(X)=n\frac{k}{N}$$

[1]-عبدلي ادريس، "**الإحصاء 03**"، مرجع سبق ذكره، ص23-24.

-التباين:

$$V(X)=n\frac{k}{N}\times\frac{N-k}{N}\times\frac{N-n}{N-1}$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{V(X)}=\sqrt{n\frac{k}{N}\times\frac{N-k}{N}\times\frac{N-n}{N-1}}$$

3-التمثيل البياني للتوزيع:

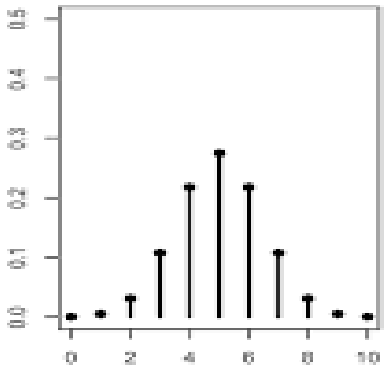
ان التوزيع فوق الهندسي يتأثر بنسبة النجاح في المجتمع (العينة)، $p=\frac{k}{N}$:

-متناظر، إذا كان الاحتمال مساويا لـ 0.5 ($p=0.5$)، هذا يعني عدد حالات النجاح وعدد حالات الفشل متساويا في المجتمع؛

-يكون التوزيع ملتويا نحو اليمين إذا كان احتمال حالة النجاح صغيرة في المجتمع، أي $p<0.5$؛

-يكون التوزيع ملتويا نحو اليسار إذا كان احتمال حالة الفشل كبيرة في المجتمع، أي $p>0.5$.

أحسن طريقة لتمثيل التوزيع فوق الهندسي بيانيا عن طريق الاعمدة المنفصلة، ليتسنى توضيح احتمال على عدد محدد من حالات النجاح، في عينة عشوائية بدون ارجاع.



4-تمارين محلولة:[1]

التمرين الأول

صندوق به عشر كرات منها ستة بيضاء وأربعة سوداء، نقوم بسحب خمس كرات دفعة واحدة، لنعتبر "X" متغير عشوائي يمثل عدد الكرات البيضاء المسحوبة.

أ-اوجد القانون الاحتمالي للمتغير العشوائي"X" ؛

ب-ما هو احتمال الحصول على ثلاث كرات بيضاء؟

ج-لنفرض اننا سحبنا ثمان كرات دفعة واحدة، اوجد القانون الاحتمالي للمتغير "X".

أ-إيجاد القانون الاحتمالي للمتغير العشوائي X:

لدينا:

$X \rightsquigarrow H(N,k,n)$

$N=10$

$k=6$

$n=5$

$X \rightsquigarrow H(10,6,5)$

$$P(X=x)=\frac{C_k^x C_{N-k}^{n-x}}{C_N^n}$$

$$P(X=x)=\frac{C_6^x C_4^{5-x}}{C_{10}^5}; x=0,1,2,3,4,5$$

ب-حساب احتمال الحصول على ثلاث كرات بيضاء:

$$P(X=3)=\frac{C_6^3 C_4^{5-3}}{C_{10}^5}$$

$$P(X=3)=\frac{C_6^3 C_4^2}{C_{10}^5}$$

$$P(X=3)=\frac{20\times 6}{252}$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

$P(X=3)=0.4761$

لو نلاحظ البسط نجد اننا سحبنا ثلاث كرات بيضاء من ست كرات بيضاء، واثنان كرات سوداء من بين أربع كرات سوداء لأننا نسحب خمس كرات.

ج-ايجاد القانون الاحتمالي للمتغير X:

عند سحبنا ثمان كرات ونحن لدينا ست كرات بيضاء وأربعة سواء، فهذا يعني ان اقل عدد للكرات البيضاء الممكنة هو أربعة، لماذا؟ لو قلنا اننا نسحب ثمان كرات من بينها صفر كرة بيضاء نجد ان هذا الحادث مستحيل، لأننا عدد الكرات السوداء هو أربع والسحب لابد ان يحتوي على ثمان كرات، لو قلنا نسحب ثمان كرات من بينها واحدة بيضاء كذاك نجد ان هذا الحادث مستحيل لأنه لابد ان يكون لدينا سبع كرات سوداء، ...، وهكذا، نجد ان اقل قيمة لعدد الكرات البيضاء المسحوبة هو أربعة (ترافقها أربع كرات سوداء)، ولا يمكن ان تتجاوز عدد الكرات البيضاء المسحوبة ست كرات لأنه يمثل أكبر عدد ممكن لها.

اذن:

$X \rightsquigarrow H(10,6,8)$

أي ان:

$$P(X=x)=\frac{C_k^x C_{N-k}^{n-x}}{C_N^n}=\frac{C_6^x C_4^{8-x}}{C_{10}^8};x=4,5,6$$

لذلك نجد في بعض المراجع عند كتابة القانون فوق الهندسي يأخذون هذه الحالة بعين الاعتبار مثلما هو موضح في العلاقة التالية:

$$P(X=x)=\frac{C_k^x C_{N-k}^{n-x}}{C_N^n};x\in\{max(0,n-N+k);\ldots;min(n,k)\}$$

التمرين الثاني

يحتوي صندوق على ثلاثة أجهزة كهربائية عاطلة وسبعة أجهزة صالحة، فاذا تم اختيار ثلاثة أجهزة وبدون إعادة وكان المتغير العشوائي "X" يمثل عدد الأجهزة العاطلة بالعينة المختارة.

أ-اكتب القيم التي من الممكن ان يأخذها هذا المتغير العشوائي واحتمال كل منها؛

ب-اوجد قيمة الاحتمالات التالية:

- $P(1<X\leq 3)$
- $P(0\leq X\leq 2)$
- $P(2\leq X<3)$
- $P(X=4)$

أ-كتابة القيم التي من الممكن ان يأخذها هذا المتغير العشوائي واحتمال كل منها:

ان:

$N=10$

$M=3$

$k=3$

$x=0, 1, 2, 3$

أي ان:

$$P(X=x)=\begin{cases}\dfrac{\binom{3}{x}\binom{7}{3-x}}{\binom{10}{3}}, & x=0,1,2,3\\ 0 & or\end{cases}$$

ومنه:

$$P(X=0)=\frac{\binom{3}{0}\binom{7}{3-0}}{\binom{10}{3}}=\frac{35}{120}$$

$$P(X=1)=\frac{\binom{3}{0}\binom{7}{3-1}}{\binom{10}{3}}=\frac{63}{120}$$

$$P(X=2)=\frac{\binom{3}{2}\binom{7}{3-2}}{\binom{10}{3}}=\frac{21}{120}$$

$$P(X=3)=\frac{\binom{3}{3}\binom{7}{3-3}}{\binom{10}{3}}=\frac{1}{120}$$

ب-ايجاد قيمة الاحتمالات التالية:

-P(1<X≤3)=P(X=2)+P(X=3)

$$P(1<X\leq 3)=\frac{21}{120}+\frac{1}{120}$$

$$P(1<X\leq 3)=\frac{22}{120}$$

$$P(1<X\leq 3)=0.1833$$

-P(0≤X≤2)=P(X=0)+P(X=1)+P(X=2)

$$P(0\leq X\leq 2)=\frac{35}{120}+\frac{63}{120}+\frac{21}{120}$$

$$P(0\leq X\leq 2)=\frac{119}{120}$$

$$P(0\leq X\leq 2)=0.9917$$

-P(2≤X<3)=P(X=2)

$$P(2\leq X<3)=\frac{21}{120}$$

$$P(2\leq X<3)=0.175$$

-P(X=4)=0

التمرين الثالث

خضع اثنا عشر مريضا بالتهاب القصبات الى تجربة طبية فاختير ستة منهم، بصورة عشوائية واعطوا العلاج "A" بينما اعطي الباقون العلاج "B".

بكم طريقة يمكن توزيع الاثنى عشر مريضا على العلاجين؟ وإذا كان أربعة من المرضى يعانون من ارتفاع ضغط الدم والآخرون طبيعيون بالنسبة لمعدل ضغط الدم، فما احتمالات:

أ-ان يخضع ذوو الضغط المرتفع جميعهم للعلاج "A"؟

ب-ان يخضع ذوو الضغط المرتفع جميعهم للعلاج نفسه؟

ج-ان يوجد واحد على الأقل من ذوي الضغط المرتفع في كل علاج؟

عدد إمكانات توزيع المرضى على العلاجين هو $\binom{12}{6}$ لان اختيار ستة وتخصيصهم للعلاج "A"، يعني بطبيعة الحال تخصيص الستة الباقين للعلاج "B".

$$\binom{12}{6}=\frac{12!}{6!6!}$$

$$\binom{12}{6}=\frac{12\times11\times10\times9\times8\times7}{6\times5\times4\times3\times2\times1}$$

$$\binom{12}{6}=924$$

أ-احتمال ان يخضع ذوو الضغط المرتفع جميعهم للعلاج "A":

لدينا:

$N=12$

$M=4$

$k=6$

$x=0, 1, 2, 3, 4$

$$P(X=x)=\frac{\binom{4}{x}\binom{8}{6-x}}{\binom{12}{6}}$$

$$P(X=4)=\frac{\binom{4}{4}\binom{8}{6-4}}{\binom{12}{6}}$$

$$P(X=4)=\frac{1}{33}$$

ب–احتمال ان يخضع ذوو الضغط المرتفع جميعهم للعلاج نفسه:

هذا يعني ان جميع ذوي الضغط المرتفع خاضعون للعلاج "A"، وجميع ذوي الضغط المرتفع خاضعون للعلاج "B".

أي:

$$P(X=4 \text{ or } 0) = P(X=0) + P(X=4)$$

$$P(X=4 \text{ or } 0) = \frac{\binom{4}{0}\binom{8}{6-0}}{\binom{12}{6}} + \frac{1}{33}$$

$$P(X=4 \text{ or } 0) = \frac{2}{33}$$

ج–احتمال ان يوجد واحد على الأقل من ذوي الضغط المرتفع في كل علاج:

هي:

$$1-\frac{2}{33}=\frac{31}{33}$$

التمرين الرابع

كتتمة للتمرين الثاني (بمعطيات التمرين الثاني)

اوجد القيمة المتوقعة والتباين للمتغير العشوائي "X".

–القيمة المتوقعة:

$$E(X)=n\frac{k}{N}$$

$$E(X)=3\frac{3}{10}$$

$$E(X)=0.90$$

–التباين:

$$V(X)=n\frac{k}{N}\times\frac{N-k}{N}\times\frac{N-n}{N-1}$$

$$V(X)=3\frac{3}{10}\times\frac{7}{10}\times\frac{7}{9}$$

$$V(X)=\frac{49}{10}$$

المحاضرة الخامسة

التوزيع الطبيعي

المحاضرة الخامسة: التوزيع الطبيعي

ان:

–التوزيع الطبيعي "*normal distribution*": يعتبر من بين اهم التوزيعات الاحتمالية ان لم نقل اهمها، ويعود ظهور هذا التوزيع الى الدراسات التي شملت كل من حركة الكواكب وأسباب وقيمة الأخطاء، وعرف أيضا بتسمية أخرى هي "منحنى قوس" "*Gaussian curve*" نسبة الى عالم الرياضيات والاحصاء الألماني "جوهان كارل فريدريش جاوس" "*Johann Carl Friedrich Gauss*" (1777-1855)، الذي وضع له الصيغة الرياضية، وقام بتطبيقه على نطاق واسع، مع بدايات القرن التاسع عشر.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع الطبيعي:

ان التوزيع الطبيعي يعد من اهم التوزيعات الاحتمالية في الإحصاء، ويرجع ذلك للأسباب الاتية: أولا ان كثيرا من الظواهر التي تظهر في التجارب العلمية تتوزع توزيعا طبيعيا، ثانيا في بعض الأحيان قد لا يكون المتغير العشوائي موزعا توزيعا طبيعيا، ولكن يمكن ان يحول الى متغير عشوائي يتوزع توزيعا طبيعيا وذلك تحت شروط معينة، ثالثا قد توجد توزيعات معقدة وبالتالي يمكن ان يستخدم التوزيع الطبيعي تقريبا لها، وأخيرا ان كافة التوزيعات الاحتمالية منفصلة كانت ام متصلة يتقارب توزيعها من التوزيع الطبيعي وذلك من خلال "نظرية النهاية المركزية" "*Central Limit Theorem*". وان هذا التوزيع يدخل في كافة المجالات الصناعية والزراعية والاقتصادية وغيرها من المجالات الأخرى.[1]

ان التوزيع الطبيعي له القدرة على تمثيل ظواهر طبيعية وبشرية متعددة وواسعة، ويمكن ان نحصر أبرز استخداماته في المجالات التالية:

-يستطيع تمثيل البيانات الحيوية والطبية، مثل درجات حرارة الجسم وضغط الدم، مما يساعد على تحديد الحالات الطبيعية والشاذة؛

-يستخدم أيضا في التعليم وعلم النفس، من اجل تمثيل توزيع علامات الاختبارات منها اختبارات الذكاء، بحيث تتمركز علامات الافراد حول المتوسط، وان ابتعدنا عنه نلاحظ التناقص التدريجي لها ما يعكس التفوق او الضعف؛

-متابعة تغيرات أسعار الأسهم والسندات لمعرفة عوائدها في الحاضر والمستقبل، من اجل تجنب مخاطرها المالية، لان النماذج الاقتصادية المصممة لذلك تعتمد على التوزيع الطبيعي لها؛

-ويستعمل لوصف الأخطاء العشوائية للقياس في ميادين الفيزياء والكمياء؛

-يستخدم في تحليل البيانات بالاعتماد على نظرية النهاية المركزية، التي تسمح تحليل احصائيات المجتمعات الكبيرة بالاستعانة بعينات صغيرة.

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص328.

1-التعريف:[1]

تعريف التوزيع الطبيعي:

يعتبر هذا التوزيع من أكثر التوزيعات الاحتمالية استخداما في النواحي التطبيقية، ومنها الاستدلال الاحصائي شاملا التقدير، واختبارات الفروض، كما ان معظم التوزيعات يمكن تقريبها الى هذا التوزيع.

إذا كان المتغير "X" متغير عشوائي له توزيع طبيعي، مداه هو "$-\infty<X<+\infty$" فان دالة كثافة احتماله هي:

$$F(X)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

حيث μ $-\infty<\mu<+\infty$ و σ^2 $(\sigma^2>0)$ يمثلان معلمتي التوزيع وهما المتوسط والتباين على التوالي.

ويرمز للتوزيع الطبيعي عادة بالرمز:

$$N\rightsquigarrow(\mu, \sigma^2)$$

ان التوزيع الطبيعي له:[2]

-منحنى ناقوسي الشكل ومتماثل وتقع قمته فوق المتوسط ويمتد طرفاه الى مالا نهاية من الجانبين، دون ان يلامسا المحور الافقي، وبالتحديد ان دالة الكثافة متماثلة حول المتوسط، وهذا يعني ان المتوسط يساوي الوسيط ويساوي المنوال، وان نقطتي الانقلاب في منحنى الدالة هما $X_1=\mu+\sigma$,$X_2=\mu-\sigma$ وانهما يقعان على بعد ثابت على يمين ويسار المنوال؛

-يتحدد بمعلومية كل من μ, σ^2 ويختلف التوزيع إذا اختلفت μ او σ^2 او كلاهما. حيث ان قيمة μ تحدد موقع التوزيع الطبيعي على الخط الحقيقي، فكلما زادت هذه القيمة تغير موقع المنحنى في الاتجاه الأيمن والعكس صحيح، بينما σ^2 تبين مقدار تشتت وتفرطح منحنى الدالة، فكلما كانت σ^2 صغيرة كلما كان المنحنى مدببا، وكلما كانت كبيرة كلما كان المنحنى متفرطح، بمعنى اخر كلما قل التباين ارتفعت قمة المنحنى وزاد تقارب الطرفين والعكس صحيح؛

-المساحة تحت منحنى الذي يمثله تساوي واحد صحيح (1)، أي ان:

---

[1]-طويطو محمد، "**مطبوعة خاصة بمقياس الإحصاء 03**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 03، 2017/2016، ص40-47.

[2]- علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص359، 361.

$$\int_{-\infty}^{+\infty} f_x(X)dx=1$$

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع الطبيعي تتمثل في:

-التوقع الرياضي:

$$E(X)=\mu$$

-التباين:

$$V(X)=\sigma^2$$

-الانحراف المعياري:

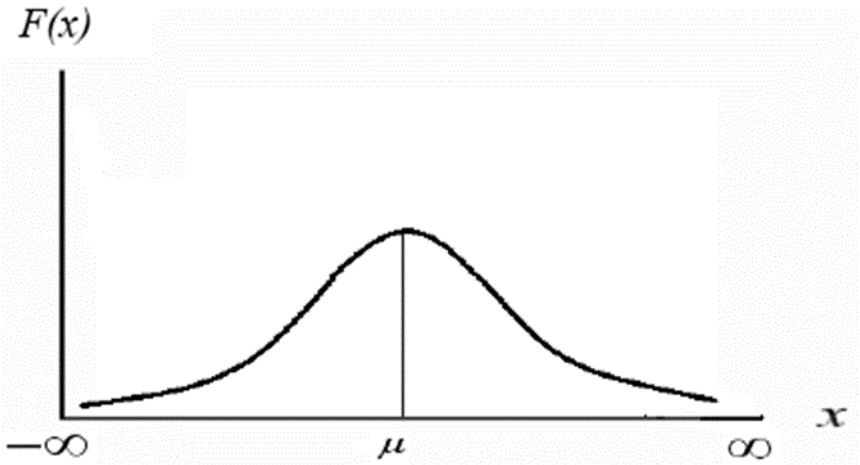
$$\sigma_X=\sqrt{V(X)}=\sqrt{\sigma^2}$$

3-التمثيل البياني للتوزيع:

يعرف التمثيل البياني للتوزيع الطبيعي بشكله الجرسي، أي ان شكل منحنى التوزيع الطبيعي هو شكل الجرس، هذا يعود الى شكله المتماثل الذي يقترب الى شكل الجرس، بحيث تتركز جل قيمه بالقرب من المتوسط (التوقع الرياضي)، وتتناقص تدريجيا كلما ابتعدنا عنه نحو الأطراف.

ان منحنى التوزيع الطبيعي يقسم الى قسمين بواسطة خط راسي يمر بالمركز الذي يمثل المتوسط والوسيط والمنوال.

ان المساحة الكلية التي تقع تحت منحنى التوزيع الطبيعي تساوي دائما الواحد الصحيح وهي عبارة عن مجموع الاحتمالات الممكنة.



4-التوزيع الطبيعي القياسي:[1]

التوزيع الطبيعي القياسي له دالة الكثافة التالية:

$$F(Z)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}Z^2}$$

من خصائص هذا التوزيع ما يلي:

-الوسط الحسابي هو:

$$E(Z)=0$$

-التباين هو:

$$V(Z)=1$$

ومن ثم يعبر عن توزيع المتغير "Z" بالرموز:

$$Z\rightsquigarrow N(0,1)$$

ويعني ذلك ان المتغير العشوائي "Z" يتبع التوزيع الطبيعي بمتوسط "0"، وتباين "1".

5-الانتقال من التوزيع الطبيعي الى التوزيع الطبيعي القياسي:[2]

لكي نستطيع الانتقال من التوزيع الطبيعي الى التوزيع الطبيعي القياسي من اجل حساب الاحتمال نتبع الخطوات التالية:

-يجب تحديد حدي المتغير واستخدام الصيغة الرياضية من اجل الانتقال من التوزيع الطبيعي الى التوزيع الطبيعي القياسي:

$$p(x_1<X<x_2) \qquad Z=\frac{(X-\mu)}{\sigma}$$

-تحويل القيم الطبيعية الى قيم طبيعية قياسية:

$$Z_1=\frac{(X_1-\mu)}{\sigma}$$

$$Z_2=\frac{(X_2-\mu)}{\sigma}$$

---

[1]-طويطو محمد، "**مطبوعة خاصة بمقياس الإحصاء 03**"، مرجع سبق ذكره، ص40-47.

[2]-نفس المرجع السابق.

-ومن ثم يكون الاحتمال:

$p(x_1<X<x_2)=p(z_1<Z<z_2)$

-تستخدم جداول التوزيع الطبيعي القياسي، والذي يعطي المساحة الخاصة بالاحتمال:

$F(z)=P(Z<z)$

6-تمارين محلولة:[1]

التمرين الأول

إذا كان الدخل السنوي للأسرة في أحد مناطق الجزائر يتبع توزيع طبيعي متوسطه 80 ألف دج، وتباينه 900.

المطلوب:

أ-كتابة قيمة معالم التوزيع للدخل السنوي؛

ب-كتابة شكل دالة كثافة الاحتمال؛

ج-ما هي نسبة الاسر التي يقل دخلها عن 60 ألف دج؟

د-ما هو الدخل الذي اقل منه 0.975 من الدخول؟

و-اوجد الاحتمالات:

- $P(Z<1.57)$
- $P(Z<-2.33)$
- $P(Z>1.96)$
- $P(-2.01<Z<1.28)$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

أ-كتابة قيمة معالم التوزيع للدخل السنوي:

بفرض ان "X" متغير عشوائي يعبر عن الدخل السنوي بألاف دج، وهو يتبع التوزيع الطبيعي، ومعالمه هي:

-المتوسط الحسابي:

$$E(X)=\mu=80$$

-التباين:

$$V(X)=\sigma^2=900$$

أي ان:

$$X\rightsquigarrow N(80,900)$$

ب-كتابة شكل دالة كثافة الاحتمال:

$$F(X)=\frac{1}{30\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-80}{30}\right)^2}$$

ج-نسبة الاسر التي يقل دخلها عن 60 ألف دج:

$$P(X<60)=P(Z<\frac{x-\mu}{\sigma})$$

$$P(X<60)=P(Z<\frac{60-80}{30})$$

$$P(X<60)=P(Z<-0.67)=F(-067)$$

وبالكشف مباشرة عن هذه القيمة في جدول التوزيع الطبيعي القياسي، نجد ان:

$$P(X<60)=P(Z<-0.67)=F(-067)=0.2514$$

د–الدخل الذي اقل منه 0.975 من الدخول:

في هذه الحالة نبحث عن قيمة المتغير "X" الذي اقل منه 0.975، بفرض ان المتغير هو "x_1"، فان:

$$P(X<x_1)=P(Z<\frac{x_1-\mu}{\sigma})$$

$$P(X<x_1)=P(Z<\frac{x_1-80}{30})=0975$$

بالكشف بطريقة عكسية، حيث نبحث عن المساحة 0.9750 نجدها تقع عند تقاطع الصف 1.9، والعمود 0.09، أي ان قيمة:

$$Z=1.96$$

ويكون:

$$\frac{x_1-80}{30}=1.96,\ Then\ x_1=30(1.96)+80$$

$$x_1=138.8$$

إذا الدخل هو 138.8 ألف دج في السنة.

و–ايجاد الاحتمالات:

-P(Z<1.57)

$P(Z<1.57)=F(1.57)=0.9418$

-P(Z<-2.33)

$P(Z<-2.33)=F(-2.33)=0.0099$

-P(Z>1.96)

$P(Z>1.96)=1-P(Z<1.96)=1-F(1.96)$

$P(Z<1.96)=0.9750$

$P(Z>1.96)=1-0.9750=0.0250$

-P(-2.01<Z<1.28)

$P(-2.01<Z<1.28)=F(1.28)-F(-2.01)$

$P(-2.01<Z<1.28)=0.8997-0.0222=0.8775$

التمرين الثاني

إذا علمت ان درجات الطلبة للدراسات العليا في امتحان القبول تتوزع وفق التوزيع الطبيعي بمتوسط 500 درجة، وانحراف معياري 100 درجة فما نسبة الطلبة الذين يتحصلون على درجة اعلى من 643 درجة؟

إذا افترضنا ان المتغير العشوائي "X" يمثل درجات الطلبة في ذلك الامتحان فان الاحتمال المطلوب هو:

$P(X>643)$

وبالتالي فان:

$$P(X>643)=P(Z>\frac{643-500}{100})$$

$$P(X>643)=P(Z>1.43)$$

$$P(X>643)=0.5-P(0<Z<1.43)$$

ومن جدول التوزيع الطبيعي نجد ان:

$$P(0<Z<1.43)=0.4236$$

وعليه فان:

$$P(X>643)=0.5-P(0<Z<1.43)$$

$$P(X>643)=0.5-0.4236$$

$$P(X>643)=0.0764$$

وذلك يعني ان 7.64% من الطلبة تحصلوا على درجة اعلى من 643 درجة، بينما 92.36% منهم تحصل على درجة اقل من 643 درجة.

التمرين الثالث

ينتج مصنع نضائد صغيرة متوسطة عمرها الزمني يساوي 76 ساعة، بانحراف معياري 10 ساعات، فاذا علمت ان العمر الزمني لهذه النضائد يتبع التوزيع الطبيعي، فأوجد احتمال ان يكون العمر الزمني لنضيدة ما:

أ-بين 71 و82 ساعة؛

ب-أكثر من 75 ساعة؛

ج-اقل من 78 ساعة.

بفرض ان المتغير العشوائي "X" يمثل العمر الزمني للنضائد الصغيرة التي ينتجها المصنع فان:

$$X \rightsquigarrow N(\mu=76, \sigma^2=100)$$

أ-حساب احتمال ان يكون العمر الزمني لنضيدة ما بين 71 و82 ساعة:

بتحويل المتغير العشوائي "X" الذي يتوزع وفق التوزيع الطبيعي الى متغير عشوائي يتبع التوزيع الطبيعي القياسي (المعياري) "Z" حتى يتسنى لنا استخدام جدول التوزيع الطبيعي المعياري، نجد ان:

$$P(71 \leq X \leq 82) = P(\frac{71-\mu}{\sigma} \leq Z \leq \frac{82-\mu}{\sigma})$$

$$P(71 \leq X \leq 82) = P(\frac{71-76}{10}) \leq Z \leq \frac{82-76}{10})$$

$$P(71 \leq X \leq 82) = P(-0.5 \leq Z \leq 0.6)$$

$$P(71 \leq X \leq 82) = P(-0.5 \leq Z \leq 0) + P(0 \leq Z \leq 0.6)$$

$$P(71 \leq X \leq 82) = 0.1915 + 0.2257$$

$$P(71 \leq X \leq 82) = 0.4172$$

ب-حساب احتمال ان يكون العمر الزمني لنضيدة ما أكثر من 75 ساعة:

$$P(X > 75) = P(Z > \frac{75-\mu}{\sigma})$$

$$P(X > 75) = P(Z > \frac{75-76}{10})$$

$$P(X > 75) = P(Z > -0.10)$$

$$P(X > 75) = P(-0.10 \leq Z \leq 0) + 0.5$$

$$P(X > 75) = 0.0398 + 0.5$$

$$P(X > 75) = 0.5398$$

ج-حساب احتمال ان يكون العمر الزمني لنضيدة ما اقل من 78 ساعة:

$$P(X<78)=P(Z<\frac{78-\mu}{\sigma})$$
$$P(X<78)=P(Z<\frac{78-76}{100})$$
$$P(X<78)=P(Z<0.2)$$
$$P(X<78)=0.5+P(0\leq Z\leq 0.2)$$
$$P(X<78)=0.5+0.0793$$
$$P(X<78)=0.5793$$

التمرين الرابع

إذا كانت درجات الحرارة في مدينة ما خلال شهر مارس تتبع التوزيع الطبيعي بمتوسط قدره $20C^o$ وانحراف معياري $3C^o$.

اوجد احتمال ان تكون درجة الحرارة بين 12 و26 درجة مئوية خلال شهر مارس.

يجب علينا إيجاد الاحتمال التالي:

$$P(12<X<26)=P(\frac{12-\mu}{\sigma}<Z<\frac{26-\mu}{\sigma})$$
$$P(12<X<26)=P(\frac{12-20}{3}<Z<\frac{26-20}{3})$$
$$P(12<X<26)=P(-2.67<Z<2)$$
$$P(12<X<26)=P(Z<2)-P(Z<-2.67)$$
$$P(12<X<26)=0.977-0.0038$$
$$P(12<X<26)=0.973$$

للوصول الى النتيجة تم الاستعانة بجدول التوزيع الطبيعي القياسي.

المحاضرة السادسة

التوزيع المنتظم

المحاضرة السادسة: التوزيع المنتظم

ان:

-التوزيع المنتظم "*Uniform distribution*": امتدت جذور ظهوره الى البدايات الأولى لتطور نظرية الاحتمالات، ولقد نسب هذا التوزيع الى عالم الرياضيات الامريكي "جايمس فيكتور اوسبنسكي" "*James Victor Uspensky*" (1883-1947)، الذي تناول مصطلح التوزيع المنتظم لأول مرة في كتابه الشهير "مقدمة في الاحتمالات الرياضية" "*Introduction to mathematical probability*"، حيث أصبحت الحاجة الى توزيع يتماشى مع تطور علم القياس، حيث يفترض ان أخطاء القياس تتبع أحيانا توزيعا منتظما ضمن نطاق جد محدد.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع المنتظم:

ويسمى في بعض الأحيان بالتوزيع المستطيل "*Rectangular distribution*" نظرا لان مخطط دالة هذا التوزيع يأخذ شكل المستطيل. ان اهم استخدامات هذا التوزيع هي تكوين ما يسمى بـ "جداول الاعداد العشوائية"، التي تستخدم في اختيار عينة عشوائية من مجتمع احصائي.[1]

نستطيع ان نعدد بعض الاستعمالات المهمة للتوزيع المنتظم، التي هي بدورها تتشابه نوعا ما مع استعمالات اغلب التوزيعات الاحتمالية الأخرى، لكن يختلف الاستعمال في كيفية الاستعمال:

-يستخدم عند تحديد نسبة التفاوت المسموح به في تصنيع المنتجات، في عملية التصنيع ان عملية التعبئة تتفاوت من وحدة الى أخرى في إطار يسمح به التصنيع (في مجال جد ضيق)، وبالتالي هذه الأخطاء مسموح بها وتتبع التوزيع المنتظم؛

-ويستخدم أيضا في تحليل المخاطر من جل التمويل، فيكون امام المحللون معلومات جد قليلة، فيكون امامهم القيمة القصوى والدنيا للأصل المالي فقط؛

-يعتمد عليه في التشفير الرقمي لبلوغ الامن السيبراني، بحيث انه في عملية التشفير يجب توفر شرط، ان تكون مفاتيح التشفير موزعة توزيع منتظم، لكي تكون عملية الاختراق جد صعبة؛

-ويستخدم في الاتصالات لمعالجة الإشارات الرقمية؛

-ويستخدم في تحليل البيانات ونتائج الألعاب مثل العاب الحظ.

[1]-نفس المرجع السابق، ص315.

1-التعريف:[1]

تعريف التوزيع المنتظم:

نقول عن المتغير العشوائي المستمر "X" انه يتبع التوزيع المنتظم إذا كانت دالة كثافته معرفة كما يلي:

$$F(X)=\begin{cases}\frac{1}{b-a}\ldots\ldots\ldots\ldots\ldots\ldots si: a\leq x\leq b\\ 0\ldots\ldots\ldots\ldots\ldots\ldots\ldots si: non\end{cases}$$

حيث ان، "a, b" مقادير ثابتة وانهما:

$a<b$

ونكتب:

$X\rightsquigarrow U(a,b)$

حتى تكون "$F(X)$" دالة كثافة احتمالية لابد من تحقق شرطين:

-لابد ان تكون:

$F(X)\geq 0$

وهذا الشرط محقق كون:

$a<b$

وبالتالي فان الفرق بينهما دوما موجب.

-ان يكون:

$$\int_{-\infty}^{+\infty} F(X)dx=1$$

وهو ما يقابل في المتغير العشوائي المنقطع بمجموع الاحتمالات لابد ان يساوي الواحد.

[1]- عبدلي ادريس، "**الإحصاء 03**"، مرجع سبق ذكره، ص30-34.

يمكن التحقق من هذا الشرط كما يلي:

$$\int_{-\infty}^{+\infty} F(X)dx = \int_{-\infty}^{a} F(X)dx + \int_{a}^{b} F(X)dx + \int_{b}^{+\infty} F(X)dx$$

$$\int_{-\infty}^{+\infty} F(X)dx = 0 + \int_{a}^{b} \frac{1}{b-a}dx + 0 = \frac{1}{b-a}\int_{a}^{b} dx = \frac{1}{b-a}[x]_{a}^{b}$$

$$\int_{-\infty}^{+\infty} F(X)dx = \frac{1}{b-a}(b-a)$$

$$\int_{-\infty}^{+\infty} F(X)dx = 1$$

ومنه "$F(X)$" هي دالة كثافة احتمالية.

اما عن دالة التوزيع لهذا المتغير العشوائي فإنها عبارة عن تكامل دالة الكثافة الاحتمالية حسب مجال تعريفها، ويكون ذلك كما يلي:

في المجال:

$$x<a$$

تكون دالة الكثافة غير معرفة وتساوي الصفر وعليه فان دالة التوزيع هي:

$$F(X) = \int_{-\infty}^{x} f(X)dx = \int_{-\infty}^{x} 0dx = 0$$

في المجال:

$$a \leq x < b$$

فان دالة التوزيع تساوي التكامل السابق نضيف اليه التكامل الذي بين أيدينا، أي الاحتمال السابق نضيف له الاحتمال الحالي وهذه الفكرة تتشابه مع المتغير العشوائي المنقطع وذلك عند البحث عن دالة التوزيع، حيث نضيف كل مرة الاحتمال السابق، للحصول في الأخير على مجموع الاحتمالات يساوي الواحد.

$$F(X) = \int_{-\infty}^{a} f(X)dx + \int_{a}^{x} f(X)dx = \int_{a}^{x} \frac{1}{b-a}dx$$

$$F(x) = \frac{1}{b-a}[x]_a^x = \frac{1}{b-a}(x-a) = \frac{x-a}{b-a}$$

في المجال:

$x \geq b$

فان دالة التوزيع هي كالاتي:

$$F(X) = \int_{-\infty}^{a} f(X)dx + \int_{a}^{b} f(X)dx + \int_{b}^{x} f(X)dx$$

وعليه فان دالة التوزيع هي كالاتي:

$$F(X) = \begin{cases} 0; si\ x < a \\ \frac{x-a}{b-a}; si\ a \leq x < b \\ 1; si\ x \geq b \end{cases}$$

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع المنتظم تتمثل في:

-التوقع الرياضي:

$$E(X) = \frac{a+b}{2}$$

-التباين:

$$V(X) = \frac{(b-a)^2}{12}$$

-الانحراف المعياري:

$$\sigma_X = \frac{b-a}{\sqrt{12}}$$

3-التمثيل البياني للتوزيع:

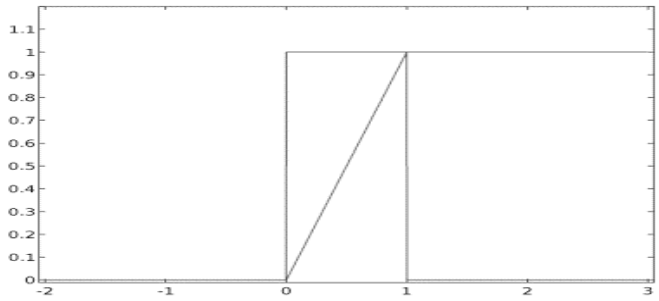
ان التمثيل البياني للتوزيع المنتظم يعرف بـ "التوزيع المستطيل"، وهذا راجع لان دالة كثافة احتماله ثابتة، أي انه عبارة عن مستطيل افقي يمتد بين القيمتين "a,b" حيث تمثل القيمة "a" القيمة الدنيا له، والقيمة "b" تمثل القيمة القصوى له. بحيث ان التمثيل البياني للتوزيع المنتظم يتم كما يلي:

-المحور الافقي، يحدد فيه القيم الممكنة لكل من "a" و"b"؛

-المحور العمودي يمثل القيمة الثابتة لدالة كثافته، وهي تساوي دائما "$\frac{1}{b-a}$"؛

-مساحة المستطيل الذي يمثل التوزيع المنتظم مساحته على الدوام تساوي الواحد (1)، وهي عبارة عن مجموع الاحتمالات.

ونستطيع ان نعطي تمثيل بياني للتوزيع المنتظم بدالتيه بالتقريب كما يلي:



6-تمارين محلولة:[1]

التمرين الأول

اذا كانت الحافلات تصل الى محطة معينة كل 15 دقيقة، ابتداء من الساعة السابعة صباحا، بمعنى تصل على تمام الساعة 07:00، 07:15، 07:30، 07:45، وهكذا، ووصل احد الركاب الى تلك المحطة في زمن يتبع التوزيع المنتظم بين الساعة السابعة والسابعة ونصف "[07:00, 07:30]"، اوجد احتمال ان الراكب سوف ينتظر الحافلة:

1-اقل من 5دقائق؛

2-أكثر من 10دقائق.

بجعل المتغير العشوائي "X" يمثل زمن وصول الراكب بعد الساعة السابعة صباحا الى تلك المحطة، نجد ان "X" يتوزع وفق التوزيع المنتظم خلال الفترة "(0، 30)".

1-وبذلك فان الراكب يجب ان ينتظر اقل من 5 دقائق إذا وفقط وصل بين الساعة 07:10 و07:15 او بين الساعة 07:25 و07:30، ويكون الاحتمال المطلوب:

$$P(10<X<15)+P(25<X<30)=\int_{10}^{15}\frac{1}{30}dx+\int_{25}^{30}\frac{1}{30}dx$$

$$P(10<X<15)+P(25<X<30)=\frac{1}{3}$$

2-وبالمثل يجب ان ينتظر أكثر من 10 دقائق إذا وصل بين الساعة 07:00 و07:05 او بين الساعة 07:15 و07:20، وبذلك يكون الاحتمال المطلوب:

$$P(0<X<5)+P(15<X<20)=\int_{0}^{5}\frac{1}{30}dx+\int_{15}^{20}\frac{1}{30}dx$$

$$P(0<X<5)+P(15<X<20)=\frac{1}{3}$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التمرين الثاني

إذا كان المتغير العشوائي "X" يتوزع وفق التوزيع المنتظم في الفترة "$[0, 10]$" فاحسب الاحتمالات التالية:

1- $P(X<3)$؛

2- $P(X>6)$؛

3- $P(3<X<8)$؛

حيث ان:

- $a=0$;
- $b=10$.

فان الصيغة العامة للتوزيع المنتظم هي:

$$f_x(X)=\begin{cases}\frac{1}{10} & , 0\leq x\leq 10\\ 0 & , \ or\end{cases}$$

وبالتالي فان:

1-

$$\mathrm{P}(X<3)=\int_0^3 \frac{1}{10}dx$$

$$\mathrm{P}(X<3)=\frac{3}{10}$$

2-

$$\mathrm{P}(X>6)=\int_6^{10} \frac{1}{10}dx$$

$$\mathrm{P}(X>6)=\frac{4}{10}$$

-3

$$P(3<X<8)=\int_3^8 \frac{1}{10}dx$$

$$P(3<X<8)=\frac{1}{2}$$

التمرين الثالث

إذا علمت ان "X" نتغير عشوائي يتوزع كتوزيع منتظم، على الفترة "$(a, -a)$" و "$a>0$" اوجد قيمة "a" بحيث ان $P(X>1)=\frac{1}{3}$

$$f(X)=\frac{1}{2a} \quad , -a<X<a$$

$$P(X>1)=\int_1^a \frac{1}{2a}dx$$

$$P(X>1)=\frac{1}{3}$$

$$\frac{1}{2a}(a-1)=\frac{1}{3}$$

ومنه فان:

a=3

التمرين الرابع

في مصنع يتم انتاج قطع غيار بشكل آلي، زمن انتاج كل قطعة يعتبر متغيرا عشوائيا موزعا بشكل منتظم بين 20 و50 دقيقة.

المطلوب:

1-استخراج دالة الكثافة الاحتمالية "$f(x)$" لزمن انتاج القطع؛

2-احسب التوقع الرياضي؛

3-احسب التباين لزمن انتاج القطعة.

1-استخراج دالة الكثافة الاحتمالية "$f(x)$" لزمن انتاج القطع:

ان زمن انتاج قطعة غيار يعتبر متغيرا عشوائيا موزعا بشكل منتظم بين 20 و50 دقيقة، يمكن استخدام هذا التوزيع لتحليل احتمالات زمن الانتظار للحصول على انتاج وحدة من قطع الغيار.

$$f(x)=\begin{cases}\frac{1}{50-20} & , \ 20\leq x\leq 50\\ 0 & , \ or\end{cases}$$

$$f(x)=\begin{cases}\frac{1}{30} & , \ 20\leq x\leq 50\\ 0 & , \ or\end{cases}$$

2-حساب التوقع الرياضي:

$$E(X)=\frac{a+b}{2}$$

$$E(X)=\frac{20+50}{2}$$

$$E(X)=35$$

3-حساب التباين لزمن انتاج القطعة:

$$\sigma^2 = \frac{(a-b)^2}{12}$$

$$\sigma^2 = \frac{(50\text{-}20)^2}{12}$$

$$\sigma^2 = \frac{(30)^2}{12}$$

$$\sigma^2 = 75$$

المحاضرة السابعة

التوزيع الأسي

المحاضرة السابعة: التوزيع الأسي

ان:

-التوزيع الأسي "*Exponential distribution*": ارتبط ظهوره بأعمال العالم الفيزياء والرياضيات السويسري "ليونهاردس أويلر" "*Leonhardus Eulerus*" (1707-1783)، الذي يعتبر الاول الذي وضع الأسس الرياضية للمتتابعات الرياضية الاسية، وأيضا الدوال الاسية، والتي تعتبر بدورها أساس التوزيع الاسي، ويعد هذا التوزيع أساسا في بعض الاعمال والنظريات مثل نظرية الطوابير، واعمال إدارة الوقت والصيانة التنبؤية.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع الأسي:

غالبا ما يستخدم هذا التوزيع في المسائل العملية الخاصة بزمن انتظار وقوع حدث معين، فمثلا الزمن الذي يبقى فيه جهاز الكتروني او الة بمصنع صالحة للعمل قبل ان يحدث بها عطب يعطلها عن العمل، او مثلا الفترة الزمنية التي ينتظرها زبون بمقهى او مصرف قبل ان تقدم اليه أي خدمة، او الفترة الزمنية ما بين وصول زبونين لمصرف او لمطعم او الى أي مكان اخر تقدم فيه خدمات، فاذا كان وقوع الحوادث يقع وفقا لنظام بواسون، فان زمن الانتظار لوقوع حدث معين والفترة الزمنية ما بين حدثين متتاليين سيكون لهما توزيع اسي. وهذه الحقيقة مفيدة جدا في استخدام التوزيع الاسي في كثير من المسائل العملية.[1]

ان أكثر استخدامات التوزيع الأسي في المجالات التي تهتم بقياس الوقت الذي يفصل بين الاحداث المستقلة وبمعدل ثابت، أهمها:

-يستخدم لحساب متوسط الوقت قبل تعطل الآلات والمعدات، وتقدير العمر الاستعمالي الافتراضي لبعض المكونات التي لا يؤثر فيها مرور الزمن، بل تتعطل فجأة؛

-يستخدم في توقع وقت الانتظار بين وصول الزبائن الى نقطة خدمة معينة؛

-ويستخدم في علوم الفيزياء النظرية لتوقع احتمال تحلل النواة في وحدة زمنية معينة ثابتة؛

-يستخدم في ترتيب وصول البيانات في شبكات التحليل لمؤسسات الاتصالات والخدمات اللوجستية؛

-يساعد على تقدير والتنبؤ بالوقت الواقع بين الحوادث والكوارث الطبيعية لأجل تقدير أقساط التامين في إدارة المخاطر والتامين.

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص354.

1-التعريف:[1]

تعريف التوزيع الأسي:

إذا كان المتغير متغيرا عشوائيا مستمرا له توزيع اسي سالب فان دالة الكثافة الاحتمالية "$F(X)$" للمتغير العشوائي المستمر "X"، تعطى وفقا للعلاقة التالية:

$$F(X)=\theta e^{-\theta x}$$

وتكون دالة الكثافة الاحتمالية معرفة كما يلي:

$$F(X)=\begin{cases}\theta e^{-\theta x} \dots\dots\dots\dots si: x>0\\ 0 \dots\dots\dots\dots si: x<0\end{cases}$$

حيث ان:

-"θ ": تمثل عدد حقيقي موجب أي ان:

$$\theta>0$$

-"e ": أساس اللوغاريتم النيبيري ويساوي:

$$e=2.71828$$

يعطى تابع التوزيع لمتغير عشوائي وليكن "X" وفقا للعلاقة التالية:

$$F(X)=P(X\leq x)$$

وبما ان "X" متغير عشوائي مستمر فان تابع التوزيع يساوي ما يلي:

$$F(X)=P(X\leq x)=\int_{-\infty}^{x} F(X)dx = 1-e^{-\theta x}$$

وعليه فان تابع التوزيع "$F(X)$" سوف يساوي ما يلي:

$$F(X)=\begin{cases}0 \dots\dots\dots\dots si: x>0\\ 1-\theta e^{-\theta x} \dots\dots\dots\dots si: 0<x<\infty\end{cases}$$

[1]-بولعباس مختار، "**محاضرات في الإحصاء 2 مدعومة بتمارين محلولة**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة تيارت، 2023/2022، ص83-85.

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع الأسي تتمثل في:

-التوقع الرياضي:

$$E(X)=\int xF(x)dx=\int \theta x e^{-\theta x}\,dx$$

لكي نستطيع حساب هذا التكامل نستعمل التكامل بالتجزئة:

$$E(X)=\frac{1}{\theta}$$

-التباين:

$$V(X)=E(x-E(X))^2=\frac{1}{\theta^2}$$

-الانحراف المعياري:

$$\sigma_X=\frac{1}{\theta}$$

3-التمثيل البياني للتوزيع:

ان الخصائص الرئيسية للتمثيل البياني للتوزيع الأسي:

-التواء التوزيع الأسي دائما يكون موجب الالتواء "*Skewed to the right*"، هذا يعني ان القيم الصغيرة في التوزيع الأسي لها أكثر تكرار؛

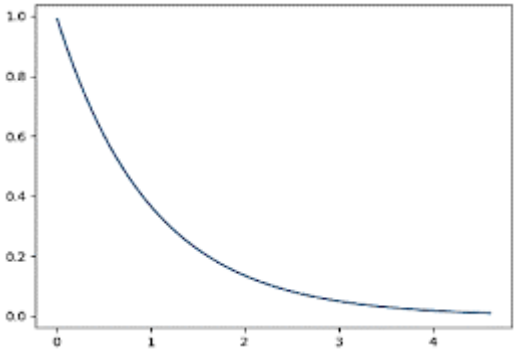
-يتميز الشكل البياني للتوزيع الاسي بالثبات عبر الوقت.

لكي نستطيع رسم الشكل البياني للتوزيع الأسي نتبع الخطوات التالية:

-نرسم المحورين، المحور الافقي (السينات) يمثل الزمن او المسافة، والمحور العمودي (العينات) ويمثل كثافة الاحتمال؛

-ان نقطة بداية التمثيل البياني للتوزيع الأسي تبدأ بأعلى نقطة في التمثيل وهي $(0, \theta)$.

ان منحنى التمثيل البياني للتوزيع الأسي يقترب من المحور الافقي (السيني) ولكن لا يلمسه ولا يقطعه، ولكن يمثل تمثيل منحنى تقاربي أي يقترب منه فقط، هذا يعني انه مع طول الوقت تصير الاحتمالات تتناقص الى اقصى درجة.



4-تمارين محلولة:[1]

التمرين الأول

ان الزمن الذي يقضيه الزبون بأحد المصارف يتبع التوزيع الأسي بمتوسط 5 دقائق. فاذا تم احتيار زبون بشكل عشوائي من ذلك المصرف فأوجد:

أ-احتمال انتظاره أكثر من 10 دقائق؛

ب-احتمال ان الزبون سينتظر من 10 دقائق علما انه انتظر أكثر من 5دقائق.

أ- احتمال انتظار زبون المصرف أكثر من 10 دقائق:

نفترض ان المتغير العشوائي "X" يمثل زمن الانتظار للزبون بالمصرف، وان:

$$\mu_X=5$$

وعليه فان:

$$\theta=\frac{1}{5}$$

$$P(X>10)=e^{-10\theta}$$

$$P(X>10)=e^{-10\frac{1}{5}}$$

$$P(X>10)=e^{-2}$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

$P(X>10)=0.1353$

ب–احتمال ان زبون المصرف سينتظر أكثر من 10 دقائق علما بانه انتظر أكثر من 5 دقائق هو:

$P(X>10/X>5)=P(X>5)$

$P(X>5)=e^{-5\frac{1}{5}}$

$P(X>5)=e^{-1}$

$P(X>5)=0.3679$

التمرين الثاني

إذا كان الزمن الذي تستغرقه مكالمة هاتفية بالدقائق بأحد الإدارات يتوزع وفق التوزيع الاسي بمتوسط 5 دقائق، وتم اختيار احدى هذه المكالمات بطريقة عشوائية فأوجد احتمال وقوع الاحداث الاتية:

أ–ان تستغرق هذه المكالمة أكثر من دقيقتين؛

ب–ان تستغرق هذه المكالمة اقل من دقيقتين؛

ج–ان تستغرق هذه المكالمة من 5 الى 10 دقائق.

أ–إيجاد احتمال ان تستغرق هذه المكالمة أكثر من دقيقتين:

$$P(X>2)=\int_2^{\infty}\frac{1}{5}e^{\frac{1}{5}X}dx$$

$P(X>2)=0.6703$

ب–إيجاد احتمال ان تستغرق هذه المكالمة اقل من دقيقتين:

$P(X<2)=1-P(X\geq 2)$

$P(X<2)=1-e^{-\frac{2}{5}}$

$P(X<2)=1-0.6703$

$P(X<2)=0.3297$

ج-إيجاد احتمال ان تستغرق هذه المكالمة من 5 الى 10 دقائق:

$$P(5\leq X\leq 10)=\int_5^{10}\frac{1}{5}e^{\frac{1}{5}X}dx$$

$P(X>2)=0.2326$

التمرين الثالث

إذا كانت اعمار أحد أنواع المصابيح الكهربائية تتبع التوزيع الاسي. وقد اختيرت اعمار مجموعة من المصابيح وشغلت لمدة 100 ساعة فوجد ان 5% منها أخفقت قبل هذه الفترة.

أ-ما النسبة المئوية من المصابيح من هذا النوع والتي يكون عمرها أطول من 200 ساعة؟

ب-إذا وجد ان مصباحا من هذا النوع قد اشتغل لمدة 100 ساعة، ما احتمال:

1-ان يستمر هذا المصباح بالعمل لما لا يقل عن 100 ساعة أخرى؟

2-ان يخفق هذا المصباح خلال الـ 100 ساعة القادمة؟

أ-إيجاد النسبة المئوية من المصابيح من هذا النوع والتي يكون عمرها أطول من 200 ساعة:

نعرف المتغير "T" ليكون عمر المصباح من هذا النوع (ساعة).

$P(T\leq 100)=0.05$

$1-e^{-100\alpha}=0.05$

$$\alpha=-\frac{ln\ (0.95)}{100}$$

$$\alpha=0.0005$$

$P(T>200)=e^{-0.0005(200)}$

$P(T>200)=0.9048$

أي ان 90.48% من المصابيح من هذا النوع يكون عمرها أطول من 200 ساعة.

ب-1-

$P(T>200/T>100)=P(T>100)$

$P(T>200)=e^{-0.0005(100)}$

$P(T>200)=0.9512$

ب-2-

$P(T\leq 200/T>100)=P(T\leq 100)$

$P(T>200)=1-e^{-0.0005(100)}$

$P(T>200)=0.0488$

التمرين الرابع

إذا كان المتغير "T" يمثل اعمار الصمامات الكهربائية "$Fuse$" التي تنتجها احدى الشركات، فاذا كان هناك طريقتان لتصنيع هذه الصمامات:

-الطريقة الأولى تعطي صمامات معدل اعمارها هو 100 ساعة؛

-الطريقة الثانية تعطي صمامات معدل اعمارها 150 ساعة.

وإذا كانت تكاليف تصنيع الصمام المنتج بطريقة الثانية يساوي ضعف تكاليف تصنيع الصمام المنتج بالطريقة الأولى. كذلك إذا كان عمر الصمام اقل من 200 ساعة فهناك غرامة قدرها "k" دينار تدفعها الشركة المنتجة عن كل صمام. بين أي الطريقتين أكثر مردودا ماديا للشركة المنتجة إذا فرضنا ان "T" يتوزع اسيا.

نفرض ان تكلفة تصنيع الصمام الواحد والمنتج بالطريقة الأولى هي "c"، فتكون تكلفة تصنيع الصمام المنتج بالطريقة الثانية هي "$2c$"، نعرف الان المتغير "Z_i" ليكون التكلفة الكلية (تكلفة التصنيع + الغرامة ان وجدت)، للصمام الواحد المنتج بالطريقة "$i=1,2$" ونلاحظ ان:

$$Z_1=\begin{cases}c & ;T>200\\ c+k & ;T\leq 200\end{cases}$$

$$E(Z_1)=cP(T>200)+(c+k)P(T\leq 200)$$

$$E(Z_1)=k(1-e^{-2})+c$$

وكذلك:

$$Z_2=\begin{cases}2c & ;T>200\\ 2c+k & ;T\leq 200\end{cases}$$

وبالنفس الطريقة السابقة:

$$E(Z_2)=k(1-e^{-\frac{4}{3}})+2c$$

$$E(Z_2)-E(Z_1)=c-0.13k$$

ونلاحظ الان التكلفة المتوقعة للطريقة الثانية تكون أكبر من تلك الطريقة الاولى إذا كانت "c" أكبر من "$0.13k$" وفي هذه الحالة فان الطريقة الأولى هي المفضلة. اما إذا كانت "c" اقل من "$0.13k$" فان الطريقة المفضلة هي الثانية.

المحاضرة الثامنة

التوزيع قاما

المحاضرة الثامنة: التوزيع قاما

ان:

–التوزيع قاما "*Gamma distribution*": يعود ظهوره الى القرن الثامن عشر على يد عالم الرياضيات السويسري "ليونهاردس أويلر" "*Leonhardus Eulerus*" (1707–1783)، الذي قام بتعريف دالة قاما الرياضية، ومع مطلع القرن التاسع عشر ساهم في تطويرها كتوزيع مستمر مجموعة من العلماء من اهمهم عالم الرياضيات الألماني "كارل فريدريش غاوس" "*Carl Friedrich Gauss*" (1777–1855)، لتستخدم بشكل واسع لنمذجة أوقات الانتظار ومشكلات المعولية، بمعدل ثابت ومتكرر.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع قاما:

ان توزيع قاما في الحقيقة مشتق من "دالة قاما" "*Gamma function*"، او ما يسمى في بعض الأحيان بـ "تكامل قاما"، الذي يرد ذكره في الكثير من كتب الرياضيات المتقدمة، ويعد هذا التوزيع كواحد من التوزيعات المهمة في دراسة المشكلات التي يكون فيها الزمن أحد عواملها، كتلك الدراسات الخاصة بطول مدة اشتغال المعدات في مصنع معين، او دراسة التعطيلات والتوقفات لماكنات مصنع معين، وكذلك يعد من التوزيعات المهمة التي تدخل في دراسة موضوع "المعولية" "*Reliability*".[1]

ان مجالات استخدام توزيع قاما هي في الاغلب في:

-فهو يعد التوزيع الأساسي للتنبؤ بفترات هطول الامطار وفترات الجفاف، ويستعمل اضا بالتنبؤ بوقوع الفيضانات من طرف مصالح الأرصاد الجوية؛

-ويستعمل في نمذجة ومعرفة العمر الإنتاجي للآلات وما تبقى من زمن لحدوث اعطال في أنظمة التشغيل، خاصة ما تعلق الامر بتراكم المشكلات مثل تآكل بعض القطع في الات الإنتاج في مصنع معين؛

-يستعمل في التنبؤ بحجم المبالغ المالية لتغطية التامين على الممتلكات، التي تكون في اغلب الأحيان نادرة الوقوع ولكن بمبالغ ضخمة؛

-ويستخدم في تمثيل نسب تراكيز الادوية في الدم، والوقت المستغرق لظهور اعراض مرض معين في حالات التعرض لعدوى؛

-يستخدم لتمثيل التداخل في إشارات الاتصال وقوة موجات الاتصال في بيئات جد معقدة.

[1]-نفس المرجع السابق، ص269.

1-التعريف:[1]

تعريف التوزيع قاما:

يستخدم توزيع قاما لتمثيل بعض الظواهر مثل توزيع الدخل والادخار تحت شروط معينة، ونقول عن متغيرة عشوائية انها تتبع التوزيع قاما إذا كانت دالة كثافتها كما يلي:

$$F(X)=\begin{cases}\dfrac{x^{\alpha-1}e^{\frac{-x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots si:x>0\\ 0\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots si:x\leq 0\end{cases}$$

$(\alpha, \beta)>0$

حيث ان:

-"$\Gamma(\alpha)$": هي دالة قاما:

$$\Gamma(\alpha)=\int_0^{\infty} t^{\alpha-1}e^{-t}dt \qquad \alpha>0$$

ونكتب:

$$X\rightsquigarrow\Gamma(\alpha,\beta)$$

[1]-جنيدي مراد، "**محاضرات في الإحصاء 3**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 3، 2021/2020، ص30-31.

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع قاما تتمثل في:

-التوقع الرياضي:

$$\mu=\alpha\beta$$

-التباين:

$$\sigma^2=\alpha\beta^2$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{\sigma^2}$$

$$\sigma_X=\sqrt{\alpha\beta^2}$$

3-التمثيل البياني للتوزيع:

يتحكم في شكل التمثيل البياني لتوزيع قاما معلمتين أساسيتين هما:

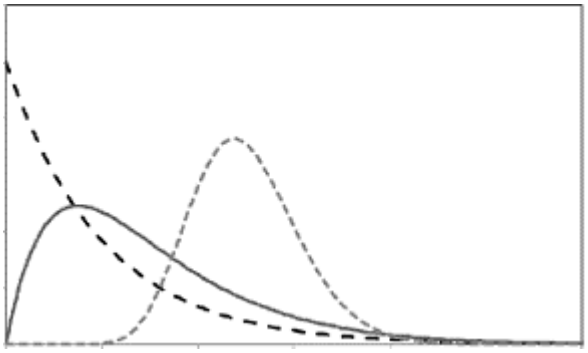
-معلمة الشكل "α" "*Shape parameter*": إذا كان:

- $\alpha=1$: فان التوزيع قاما ينتقل الى التوزيع الأسي، فيبدأ التمثيل البياني لتوزيع قاما من الأعلى ثم ينحدر الى الأسفل؛

- $\alpha<1$: فان التوزيع قاما يكون تمثيله البياني شديد الالتواء ويقترب من ما لا نهاية عند القيمة الصفر؛

- $\alpha>1$: فان التوزيع قاما يكون تمثيله البياني مشابها لشكل الجرس ولكن ملتويا جهة اليمين "*Right Skewed*"، وكلما زادت قيمة α يصبح شكل التوزيع قاما أكثر تماثلا ويقترب الى شكل التوزيع الطبيعي.

-معلمة القياس "β" "*Scale parameter*": ان هذه المعلمة تتحكم في مدى انتشار او تسطح التمثيل البياني لتوزيع قاما، فكلما زادت β كل زاد تسطح التمثيل البياني لتوزيع قاما.



4-تمارين محلولة:[1]

التمرين الأول

احسب المتوسط والتباين للمتغيرات العشوائية التالية:

$$F(X)=\begin{cases}\dfrac{x^4e^{\frac{-x}{2}}}{e^5\Gamma(5)} & x>0\\ 0 & x\leq 0\end{cases} \qquad F(Y)=\begin{cases}\dfrac{y^3e^{\frac{-y}{4}}}{4^4(6)} & y>0\\ 0 & y\leq 0\end{cases} \qquad F(Z)=\begin{cases}\dfrac{z^2e^{-z}}{6} & z>0\\ 0 & z\leq 0\end{cases}$$

-$\mu_X=\alpha\beta=5\times 2$

$\mu_X=10$

$\sigma_X^2=\alpha\beta^2=5\times 2^2$

$\sigma_X^2=20$

-$\mu_Y=\alpha\beta=4\times 4$

$\mu_Y=16$

$\sigma_Y^2=\alpha\beta^2=4\times 4^2$

$\sigma_Y^2=64$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

$-\mu_Z=\alpha\beta=3\times1$

$\mu_Z=3$

$\sigma_Z^2=\alpha\beta^2=3\times1^2$

$\sigma_Z^2=3$

التمرين الثاني

إذا كان المتغير العشوائي "X" يمثل اعمار الانابيب الالكترونية (ساعة)، التي تنتجها احدى الشركات وكانت:

$F(X)=\alpha^2Xe^{-2\alpha X} \quad ;X>0$

أ-إيجاد توقع ان معدل اعمار الانابيب المنتجة هو $\frac{1}{\alpha}$ ساعة؛

ب-ايجاد احتمال ان لا يقل عمر انبوبة عن "m" هو:

$P(X>m)=(1+\alpha m)e^{-\alpha m}$

نلاحظ ان "X" يتبع توزيع قاما بالمعلمتين $\alpha, 2$.

أ-إيجاد توقع ان معدل اعمار الانابيب المنتجة هو $\frac{1}{\alpha}$ ساعة:

اذن:

$$E(X)=\frac{2}{\alpha}$$

ب-ايجاد احتمال ان لا يقل عمر انبوبة عن "m":

وعلى ما سبق فان:

$P(X>m)=1-P(X\leq m)$

$P(X>m)=1-F(m)$

$$P(X>m)=\sum_{k=0}^{2-1}\frac{e^{-\alpha m}(\alpha m)^k}{k!}$$

$P(X>m)=e^{-\alpha m}+e^{-\alpha m}(\alpha m)$

$P(X>m)=e^{-\alpha m}(1+\alpha m)$

التمرين الثالث

احسب التوقع والتباين للمتغيرات العشوائية التالية:

$$F(X)=\begin{cases}\dfrac{x^4e^{-\frac{x}{2}}}{2^5\Gamma(5)} & ,x>0\\ 0 & ,x\leq 0\end{cases}\quad F(Y)=\begin{cases}\dfrac{y^3e^{-\frac{y}{4}}}{4^4(6)} & ,y>0\\ 0 & ,y\leq 0\end{cases}\quad F(Z)=\begin{cases}\dfrac{z^2e^{-z}}{6} & ,z>0\\ 0 & ,z\leq 0\end{cases}$$

$$F(X)=\begin{cases}\dfrac{x^4e^{-\frac{x}{2}}}{2^5\Gamma(5)} & ,x>0\\ 0 & ,x\leq 0\end{cases}$$

-التوقع الرياضي:

$\mu=\alpha\beta$

$\mu=5(2)$

$\mu=10$

-التباين:

$\sigma^2=\alpha\beta^2$

$\sigma^2=5(2)^2$

$\sigma^2=20$

$$F(Y)=\begin{cases}\dfrac{y^3e^{-\frac{y}{4}}}{4^4(6)} & ,y>0\\ 0 & ,y\leq 0\end{cases}$$

-التوقع الرياضي:

$\mu=\alpha\beta$

$\mu=4(4)$

$\mu=16$

-التباين:

$\sigma^2=\alpha\beta^2$

$\sigma^2=4(4)^2$

$\sigma^2=64$

$$F(Z)=\begin{cases}\frac{z^2e^{-z}}{6} & ,z>0\\ 0 & ,z\leq 0\end{cases}$$

-التوقع الرياضي:

$\mu=\alpha\beta$

$\mu=3(1)$

$\mu=3$

-التباين:

$\sigma^2=\alpha\beta^2$

$\sigma^2=3(1)^2$

$\sigma^2=3$

التمرين الرابع

إذا كانت اعمار منتجات مصنع معين "X" بالسنين تتبع التوزيع الاحتمالي الاتي:

$F(X)=cXe^{-2X}$ $;X\geq 0$; $c=const$

أ-اوجد قيمة "c"؛

ب-إذا اخترنا عشوائيا وحدة ما من هذا الإنتاج، فما احتمال ان يعيش اقل من 3 شهور؟

يتضح مما سبق ان "X" يتبع التوزيع قاما بالمعلمات التالية:

$n=2$
$\lambda=2$

أ-ايجاد قيمة "c":

وعلى ذلك فان:

$$c=\frac{\lambda^n}{\Gamma(n)}$$
$$c=\frac{2^2}{\Gamma(2)}$$
$$c=4$$

ب-إيجاد احتمال ان يعيش اقل من 3 شهور:

ان ثلاثة أشهر يعني:

$$3\ months=\frac{1}{4}\ year$$
$$3\ months=0.25\,year$$

اذن:

$$P(X\leq 0.25)=4\int_0^{0.25} Xe^{-2X}dX$$
$$P(X\leq 0.25)=[-2Xe^{-2X}-e^{-2X}]_0^{0.25}$$
$$P(X\leq 0.25)=0.09$$

المحاضرة التاسعة

التوزيع بيتا

المحاضرة التاسعة: التوزيع بيتا

ان:

–التوزيع بيتا "*Beta distribution*": ارتبط ظهوره عبر التاريخ بمشكلة "الاستدلال البايزي"، ومن أوائل من استخدموا شكل من اشكال هذا التوزيع العالمين "توماس بايز" (1701-1761) و"ريتشارد برايس" (1723-1791) "*Thomas Bayes and Richard Price*"، لأجل وصف نجاح حدث غير محدد او غير معروف، وفي عام 1774 تم توسيع أفكار "بايز" واستخدامات توزيع بيتا من طرف العالم الرياضي الفرنسي "بيير سيمون لابلاس" (1749-1827) "*Pierre Simon Laplace*"، لتقدير احتمالات وقوع الاحداث المستقبلية، بناء على المعلومات السابقة، هذه الأفكار أدت الى ظهور "قاعدة لابلاس للتتابع". وأول من قام بصياغة "تكامل بيتا" رياضيا العالم السويسري "ليونهاردس أويلر" "*Leonhardus Eulerus*" (1707-1783) في القرن الثامن عشر، ولكن العالم "أدريان ماري ليجاندر" "*Adrien-Marie Legendre*" (1752-1833) هو من أعطاه تسمية توزيع بيتا.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع بيتا:

ان هذا التوزيع مشتق من دالة بيتا "*Beta function*" او ما تسمى في بعض الأحيان تكامل بيتا، ويعد واحد من التوزيعات ذات أهمية تطبيقية في حقل الرقابة على جودة الإنتاج من خلال تكوين ما يسمى "جداول عينات القبول" "*Acceptance sampling tables*"، التي تستخدم في اتخاذ القرار بشأن قبول وحدات الإنتاج استنادا الى نسب الوحدات المعيبة في العينة.[1]

ويستخدم توزيع بيتا:

-لإدارة المشاريع خصوصا بتطبيق طريقة "*Pert*"، في تقدير أوقات الأنشطة اللازمة لإنجاز المشاريع؛

-لتحديد النسب في بعض المجالات، مثل تحديد نسبة الرطوبة او درجات الحرارة لمنطقة او تربة ما، او لنسبة الاستجابة عند المرضى لدواء معين؛

-ليساعد على تحديد التصميمات الأكثر نجاحا للمواقع الالكترونية التي تستخدم في التسويق الالكتروني للمنتجات؛

-بشكل أوسع لمراقبة الجودة، يستخدم لمعرفة الوحدات التي تحمل عيوب في عملية مراقبة الإنتاجية في المصانع، او عمليات التفتيش للشحنات المنتجة من قبل.

[1]-نفس المرجع السابق، ص385.

1-التعريف:[1]

تعريف التوزيع بيتا:

يتميز توزيع بيتا بمرونة الكبيرة تبعا لقيم معلمتيه حيث يستخدم لحساب توزيع "t^2" و توزيع "F"، التوزيع الثنائي، الثنائي السالب وغيرها، وتستخدم لتمثيل بعض المتغيرات التي تتراوح بين 0 و1، مثل نسبة ما، كنسبة التلف او المبيعات، ... .

نقول عن متغيرة عشوائية انها تتبع توزيع بيتا إذا كانت دالة كثافتها كما يلي:

$$F(X)=\begin{cases}\dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} & \ldots\ldots\ldots\ldots si: 0<x<1\\ 0 & \ldots\ldots\ldots\ldots si: non\end{cases}$$

$(\alpha, \beta)>0$

حيث ان:

"$B(\alpha,\beta)$": هي دالة بيتا

$$B(\alpha,\beta)=\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du \quad (\alpha,\beta)>0$$

ونكتب:

$$X\rightsquigarrow B(\alpha,\beta)$$

والصيغة الرياضية التي تربط كل من توزيع قاما بتوزيع بيتا هي:

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

[1]-جنيدي مراد، "**محاضرات في الإحصاء 3**"، مرجع سبق ذكره، ص31-32.

2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع بيتا تتمثل في:

-التوقع الرياضي:

$$\mu=\frac{\alpha}{\alpha+\beta}$$

-التباين:

$$\sigma^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

-الانحراف المعياري:

$$\sigma_X=\sqrt{\sigma^2}$$

$$\sigma_X=\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}$$

3-التمثيل البياني للتوزيع:

ان التمثيل البياني لتوزيع بيتا يأخذ عدة اشكال، وهو محصور بين الصفر (0) والواحد (1)، وشكل التمثيل البياني لتوزيع بيتا يحدد عن طريق المعلمتين "α, β ":

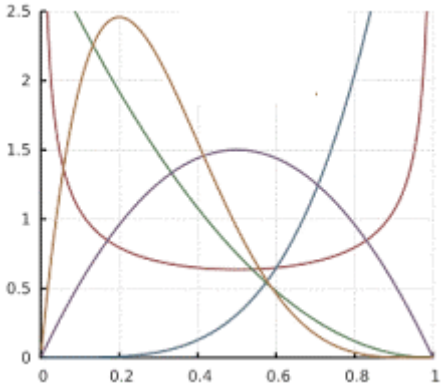
-شكل توزيع بيتا متناظر (متماثل) "*Symmetric*": يأخذ شكل الجرس مركزه 0.5، إذا كانت $\alpha=\beta$ وتكون كليهما أكبر من 1؛

-شكل توزيع بيتا موحد "*Uniform*": عندما $\alpha=\beta=1$ فيكون التمثيل البياني لتوزيع بيتا عبارة عن خط افقي؛

-شكل توزيع بيتا له التواء نحو اليمين "*Right-Skewed*": عندما تكون $\alpha<\beta$؛

-شكل توزيع بيتا له التواء نحو اليسار "*Left-Skewed*": عندما تكون $\alpha>\beta$؛

-شكل توزيع بيتا حسب الحرف اللاتيني "U": عندما تكون كل من المعلمتين α, β اقل من 1، فيرتفع عند طرفيه وينخفض عند المنتصف.



4-تمارين محلولة:[1]

التمرين الأول

احسب ما يلي:

$B(3,4)$ $B\left(\frac{1}{2},\frac{1}{2}\right)$ $B(n,2)$ $n \in N$

- $\mathbf{B(3,4)=\frac{(3-1)!(4-1)!}{(7-1)!}}$

$$B(3,4)=\frac{2}{120}$$

$$B(3,4)=\frac{1}{60}$$

- $\mathbf{B\left(\frac{1}{2},\frac{1}{2}\right)=\frac{(\sqrt{\pi})(\sqrt{\pi})}{\left(\frac{1}{2}+\frac{1}{2}\right)}}$

$$B\left(\frac{1}{2},\frac{1}{2}\right)=\pi$$

- $\mathbf{B(n,2)=\frac{(n-1)!(1)!}{(2+1)!}}$

$$B(n,2)=\frac{1}{n(n+1)!}$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التمرين الثاني

إذا كان لدينا المتغير العشوائي المستمر "X" الذي يمثل نسبة الرطوبة قي مدينة الجزائر، وكانت دالة التوزيع الاحتمالي لذلك كما يلي:

$$F(X)=\begin{cases}12x^2(1-x) ; & 0<x<1\\ 0 ; & or\end{cases}$$

أ-ما هو احتمال ان تكون نسبة الرطوبة في المدينة على الأكثر 30%؟

ب-ما هو احتمال ان تكون نسبة الرطوبة في المدينة على الأقل 20%؟

ج-احسب التوقع والتباين لنسبة الرطوبة في هذه المدينة.

أ-حساب احتمال ان تكون نسبة الرطوبة في المدينة على الأكثر 30%:

$$P(X\leq 0.3)=\int_0^{0.3} 12x^2(1-x)dx$$

$$P(X\leq 0.3)=0.0837$$

$$P(X\leq 0.3)=8.38\%$$

ب-حساب احتمال ان تكون نسبة الرطوبة في المدينة على الأقل 20%:

$$P(X\geq 0.2)=\int_{0.2}^{1} 12x^2(1-x)dx$$

$$P(X\leq 0.3)=0.9728$$

$$P(X\leq 0.3)=97.28\%$$

ج-حساب التوقع والتباين لنسبة الرطوبة في هذه المدينة:

-التوقع:

$$\mu=\frac{\alpha}{\alpha+\beta}$$

$$\mu=\frac{3}{3+2}$$

$$\mu=0.6$$

-التباين:

$$\sigma^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\sigma^2=\frac{3(2)}{(5)^2(6)}$$

$$\sigma^2=0.06$$

التمرين الثالث

ان نسبة استهلاك ذاكرة تخزين السيرفر "*Server*" في احدى الشركات تتبع التوزيع بيتا بالمعلمات التالية:

$\alpha=2$
$\beta=1$

أ-اوجد دالة الكثافة الاحتمالية $F(X)$؛

ب-احسب احتمال ان يكون استهلاك ذاكرة تخزين السيرفر 80%.

أ-اوجد دالة الكثافة الاحتمالية *F(X)*:

$$F(X)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

$$B(\alpha,\beta)=\frac{\Gamma(2)\Gamma(1)}{\Gamma(2+1)}$$

$$B(\alpha,\beta)=\frac{1!\times 0!}{2!}$$

$$B(\alpha,\beta)=\frac{1}{2}$$

$$F(X)=\frac{1}{1/2}x^{2-1}(1-x)^{1-1}$$

$$F(X)=2x$$

ب-احسب احتمال ان يكون استهلاك ذاكرة تخزين السيرفر 80%:

$$P(X>0.8)=\int_{0.8}^{1} 2xdx$$

$$P(X>0.8)=[x^2]_{0.8}^{1}$$

$$P(X>0.8)=(1)^2-(0.8)^2$$

$$P(X>0.8)=0.36$$

ان احتمال ان يتجاوز استهلاك ذاكرة تخزين السيرفر 80% هو 36%.

التمرين الرابع

نفترض ان نسبة النجاح في امتحان معين تتبع توزيع بيتا بالمعلمتين:

$\alpha=8$
$\beta=2$

احسب كل من التوقع والتباين لنسبة النجاح في هذا الامتحان.

-التوقع:

$$\mu=\frac{\alpha}{\alpha+\beta}$$

$$\mu=\frac{8}{8+2}$$

$$\mu=\frac{8}{10}$$

$$\mu=0.8$$

-التباين:

$$\sigma^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$\sigma^2=\frac{8(2)}{(10)^2(11)}$$

$$\sigma^2=\frac{16}{1100}$$

$$\sigma^2=0.0145$$

المحاضرة العاشرة

التوزيع كاي مربع

المحاضرة العاشرة: التوزيع كاي مربع

ان:

–التوزيع كاي مربع "*Chi-square distribution*": يعود تاريخ اول ظهور له الى عالم الإحصاء الانجليزي الشهير "كارل بيرسون" "*Karl Pearson*" (1857–1936) مع مطلع القرن العشرين، بالضبط مع بدايات 1900، ويستخدم هذا التوزيع لتقييم ما إذا كان هناك تطابق بين البيانات المجمعة والتوقعات المحددة مسبقا، أي يستخدم بصورة كبيرة في اختبار قبول او رفض الفرضيات، بالاعتماد على الجداول الاحتمالية الخاصة به. ويستمد اسمه من الحرف اليوناني "خي" "*Chi*" الذي قام باستعماله "كارل بيرسون" عند صياغته لهذا التوزيع رياضيا.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

-التوزيع كاي مربع:

يعتبر التوزيع التكراري الاحتمالي كاي مربع من التوزيعات التكرارية المتصلة، والذي يستخدم بكثافة للإجابة عن الكثير من التساؤلات حول المجتمع الاحصائي، مثل التوافق بين مجاميع معينة من المعلومات "*Goodnesss of fit*" او تجانس بين مجموعة من العينات أي ما إذا كانت مأخوذة او تابعة لنفس المجتمع الاحصائي.[1]

ان أبرز استخدامات التوزيع كاي مربع هي:

-يستخدم لتحديد اختبارات جودة المطابقة "*Goodness-of-Fit Tests*"، إذا كانت عينة البيانات التي سحبت تتوافق مع بعض التوزيعات الاحتمالية المعروفة؛

-يستخدم أيضا في اختبار الاستقلالية "*Test of Independence*" بين متغيرتين كيفيتين (نوعيتين)، وبالتالي يستخدم لفحص هل يوجد علاقة من عدم وجود العلاقة بين متغيرتين كيفيتين؛

-لاختبار تطابق سمة او عدة سمات بين مجموعات سكانية مختلفة، أي يستخدم لاختبارات التجانس بين التوزيعات "*Test of Homogeneity*"؛

-عندما يكون التوزيع طبيعي يستعمل لتقدير الانحراف المعياري وبالتالي تباين المجتمع في الإحصاء المعلمي "*Parametric statistics*"؛

-في حالة البيانات الضخمة يستعمل من اجل تقييم جودة النماذج التنبؤية باختبار المعنوية لمعاملات نماذج الانحدار اللوجيستي "*Logistic Regression Coefficients*".

[1]-علاء الدين القبانجي وحسام حمامة كمرجي، "**الاحتمال والاحصاء**"، منشورات جامعة دمشق، 2012، ص195.

1-التعريف:[1]

تعريف توزيع كاي مربع:

يعتبر توزيع كاي مربع من بين اهم التوزيعات الاحتمالية المستمرة، ويستعمل في اختبار الفرضيات والمقارنة بين عينتين، كما يستعمل في تحديد الاستقلالية بين حدثين.

إذا كان لدينا "n" متغيرة عشوائية "$X_1, X_2, \ldots, X_n$" مستقلة فيما بينها، وتخضع للتوزيع الطبيعي المعيلري:

$$N(1,0)$$

فان:

$$X_1^2+X_1^2+\ldots+X_1^2$$

تتبع توزيع كاي مربع بدرجة حرية "n" ويرمز له بالرمز:

$$\chi^2$$

$$\chi^2 = X_1^2+X_1^2+\ldots+X_1^2 = \sum_{i=1}^{n} \chi_i^2$$

تعرف دالة كثافة احتمال التوزيع كاي مربع على الشكل التالي:

$$F(X)=\begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} X^{(\frac{n}{2}-1)e^{-\frac{X}{2}}} & ;X>0 \\ 0 & ;X\leq 0 \end{cases}$$

حيث ان "Γ" هي دالة قاما وتحسب بالعلاقة التالية:

$$\Gamma=\int_0^{+\infty} X^{n-1}e^{-X}dx=(n-1)!$$

[1]-همال فريدة، "**الإحصاء 3 دروس وتمارين**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 3، 2022/2021، ص10-12.

تعرف دالة التوزيع التراكمية لتوزيع كاي مربع على الشكل التالي:

$$F(X)=P(X^2\leq x)=P(T\leq t)=\int_0^x \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} t^{(\frac{n}{2}-1)} e^{-\frac{t}{2}} dt$$

تستخرج الاحتمالات والقيم الحرجة من الجدول الاحصائي لتوزيع كاي مربع، حيث الجدول يعطينا مساحة الاحتمال في حالة:

$$P(\chi^2\geq x)$$

وما تبقى من الحالات تحسب من العلاقات التالية:

$$P(\chi^2\leq x)=1-P(\chi^2\geq x)$$

$$P(x_1\leq\chi^2\leq x_2)=P(\chi^2\geq x_1)-P(\chi^2\geq x_2)$$

2-المميزات العددية للتوزيع:

بالاستعانة بقانون توزيع قاما فان المميزات العددية للتوزيع كاي مربع تتمثل في:

-التوقع الرياضي:

$$E(\chi^2)=\frac{\frac{n}{2}}{\frac{1}{2}}$$

$$E(\chi^2)=n$$

-التباين:

$$V(\chi^2)=\frac{\frac{n}{2}}{\frac{1}{4}}$$

$$V(\chi^2)=2n$$

-الانحراف المعياري:

$$\sigma_{\chi^2}=\sqrt{V(\chi^2)}$$

3-التمثيل البياني للتوزيع:

يتميز التمثيل البياني لتوزيع كاي مربع بما يلي:

-ان التمثيل البياني لتوزيع كاي مربع يبدا دائما من نقطة الصفر ويمتد الى ما لا نهاية جهة اليمين، هذا ما يجعل شكله ملتويا جهة اليمين "*Positively Skewed*"؛

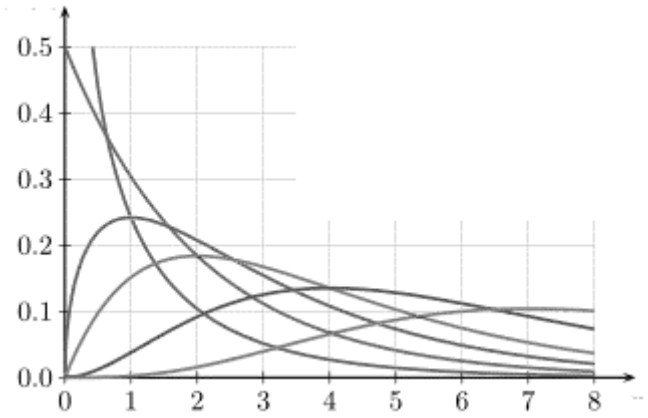
-ان التمثيل البياني لتوزيع كاي مربع يتأثر بدرجة كبيرة بدرجات الحرية حيث ان:

-عندما تأخذ درجة الحرية القيمتين 1 او 2 فان المنحنى يأخذ اعلى قيمة له عند الصفر ثم يبدا بالانحدار للأسفل؛

-يأخذ شكل الجرس بزيادة درجات الحرية وبالتالي تنتقل قمته جهة اليمين؛

-يقترب التمثيل البياني لتوزيع كاي مربع من التمثيل البياني للتوزيع الطبيعي إذا كانت درجة الحرية كبيرة جدا.

-يتميز أيضا التمثيل البياني لتوزيع كاي مربع بان جميع قيمه موجبة ولا يكون معرف على القيم السالبة وهذا راجع الى التربيع في صيغته الرياضية، وبالتالي يقع بأكمله في جهة اليمين من المحور الافقي، أي من الجهة الموجبة فقط.



4-تمارين محلولة:[1]

التمرين الأول

قم بإيجاد الاحتمالات لكل من:

- $P(\chi^2_{13}<9.929)$؛
- $P(\chi^2_{26}<-10)$؛
- $P(\chi^2_{25}<34.382)$؛
- $P(9.591\leq\chi^2_{20}\leq34.178)$.

$-P(\chi^2_{13}<9.929)=1-P(\chi^2_{13}\geq9.929)$

$P(\chi^2_{13}<9.929)=1-0.70$

$\boldsymbol{P(\chi^2_{13}<9.929)=0.30}$

$\boldsymbol{-P(\chi^2_{26}<-10)=0}$

$\boldsymbol{-P(\chi^2_{25}<34.382)=0.10}$

$-P(9.591\leq\chi^2_{20}\leq34.178)=P(\chi^2_{20}\geq9.591)+P(\chi^2_{20}\geq34.178)$

$P(9.591\leq\chi^2_{20}\leq34.178)=0.90-0.035$

$\boldsymbol{P(9.591\leq\chi^2_{20}\leq34.178)=0.865}$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التمرين الثاني

إذا كان لدينا الاحتمال يقدر بـ 0.95، اوجد:

-القيمة الحرجة لهذا الاحتمال عند درجة حرية n=6؛

-اوجد $P(\chi^2>15.1)$ إذا كان:

$\chi^2_{p,n}=\chi^2_{0.05,5}=15.1$

تستخرج كل من القيمة الحرجة والاحتمال من الجدول الاحصائي لتوزيع كاي مربع:

$\chi^2_{0.95,6}=12.6$

$P(\chi^2>15.1)=1-P(\chi^2>15.1)$

$P(\chi^2>15.1)=1-0.99$

$P(\chi^2>15.1)=0.01$

التمرين الثالث

ليكن "X" متغير عشوائي يتبع التوزيع كاي مربع بدرجة حرية 3 أي ان:

$X\rightsquigarrow\chi^2_3$

أ-اكتب دالة كثافة التوزيع؛

ب-احسب كل من التوقع الرياضي والتباين؛

ج-اوجد قيمة X التي توافق الاحتمال $P(X\geq\chi^2_3)=0.05$.

أ-كتابة دالة كثافة التوزيع:

$$F(X)=\frac{X^{(\frac{1}{2})}e^{-\frac{X}{2}}}{2^{\frac{3}{2}}\Gamma(\frac{3}{2})}$$

ب–حساب كل من التوقع الرياضي والتباين:

-التوقع الرياضي:

$E(\chi^2)=n$

$E(\chi^2)=3$

–التباين:

$V(\chi^2)=2n$

$V(\chi^2)=2(3)$

$V(\chi^2)=6$

ج–ايجاد قيمة X التي توافق الاحتمال $P(X\geq\chi_3^2)=0.05$:

بواسطة الجدول الاحصائي لتوزيع كاي مربع فان القيمة الحرجة للاحتمال:

$P(X\geq\chi_3^2)=0.05$

$X=7.81$

التمرين الرابع

إذا كانت "χ_n^2" ترمز لمتغير عشوائي يتبع توزيع كاي بدرجات حرية تساوي "n"، من جدول كاي مربع اوجد:

$-\chi^2_{0.01,5}$؛

$-\chi^2_{0.05,20}$؛

$-\chi^2_{0.90,28}$؛

بالاستعانة بالجدول الاحتمالي للتوزيع كاي مربع، فان:

$-\chi^2_{0.01,5}$

$n=5$;

$\alpha=0.01$

$\chi^2_{0.01,5}=15.070$

$-\chi^2_{0.05,20}$

$n=20$;

$\alpha=0.05$

$\chi^2_{0.05,20}=31.410$

$-\chi^2_{0.90,28}$

$n=28$;

$\alpha=0.90$

$\chi^2_{0.90,28}=18.939$

المحاضرة الحادية عشر

توزيع ستيودنت وتوزيع فيشر

المحاضرة الحادية عشر: توزيع ستيودنت وتوزيع فيشر

ان:

–التوزيع ستيودنت "*Student's t-distribution*": من التوزيعات الاحتمالية الجد مهمة، وسبب ظهوره كان تصحيح الاختلال في التوزيع الطبيعي عندما كان يستخدم في العينات الصغيرة، التي تقل في العادة عن 30 مفردة، او عندا لا نعرف معلمات المجتمع مثل تباين او الانحراف المعياري، وأول من استعمل تسمية وتوزيع ستيودنت "*Student's*" في أوائل القرن العشرين هو العالم الاحصائي والكيميائي الانجليزي "ويليام سيلي جوست" "*William Sealy Gosset*" (1876–1937)، لم ينسب التوزيع له بالاسم لأنه كان ممنوعا عليه النشر باسمه؛

–التوزيع فيشر "*Fisher distribution*": يعود ظهوره لعالم الإحصاء وعلم الاحياء الانجليزي "رونالد آيلمر فيشر" "*Ronald Aylmer Fisher*" (1890–1962)، الذي لعب الدور الرئيسي في تأسيس علم الوراثة السكانية، فاخذ تسميته، ولقد تطور عبر مرور الزمن بإضافات مهمة لمجموعة من العلماء، ويعتبر هذا التوزيع أداة جد مهمة لتقييم الفروق بين البيانات.

المتغير العشوائي المستمر (المتصل) "*Continuous random variable*":

يقال "X" متغير عشوائي مستمر إذا كان فضاء العينة مجموعة غير قابلة للعد، سواء كانت منتهية ام غير منتهية، او بمعنى اخر فان اية دالة ذات قيمة حقيقية معرفة على فضاء عينة مستمر تسمى متغيرا عشوائيا مستمرا.[1]

[1]–أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص41.

1-التوزيع ستيودنت:

ان التوزيع ستيودنت يطلق عليه "توزيع t"، وهو ذو علاقة بالعينات العشوائية التي يتم اختيارها من مجتمع طبيعي، ولهذا التوزيع تطبيقات عديدة في مجال الإحصاء الاستنتاجي، ولقد وضعت جداول خاصة به ويعرف هذا التوزيع كما يلي، إذا كان "X" متغيرا عشوائيا يتبع التوزيع الطبيعي المعياري "$N(0,1)$"، وكان "Y" متغيرا عشوائيا له توزيع كاي مربع، بدرجة حرية "n" وكان المتغير العشوائي "Y" مستقلا عن المتغير العشوائي "X"، فان المتغير العشوائي "$T=\frac{X}{\sqrt{\frac{Y}{n}}}$" له توزيع يطلق عليه تسمية "توزيع t"، بدرجات حرية تساوي "n".[1]

ان الاستخدامات الأساسية لتوزيع ستيودنت هي:

-كاختبار للمقارنة بين الفروق المعنوية بين معلمات المجموعات التي تكون في الغالب حجمها صغير؛

-لتقدير فترة الثقة التي يقع فيها متوسط المجتمع بناء على بيانات قليلة؛

-لاختبار مدى معنوية معاملات العلاقات الانحدارية بين المتغيرات أي العلاقة الارتباطية بين المتغيرات؛

-في عمليات الإنتاج من اجل مراقبة المنتوج من خلال سحب عينة عشوائية تكون صغيرة الحجم، بفحص مدى مطابقتها لمعايير الإنتاج المعتمدة.

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص391.

1-1-التعريف:[1]

تعريف توزيع ستيودنت:

لتكن المتغيرتان العشوائيتان المستقلتان "Z, Y" حيث ان:

$$Y \rightsquigarrow N(0,1)$$

و:

$$Y \rightsquigarrow \chi_v^2$$

فان المتغيرة العشوائية "T":

$$T=\frac{Y}{\sqrt{\frac{Z}{v}}}$$

تتبع التوزيع ستيودنت.

ان دالة كثافة المتغيرة العشوائية "T" تكتب على الشكل الموالي:

$$f(T)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}$$

ونقول ان المتغيرة "T" تتبع التوزيع ستيودنت بـ "v" درجة حرية ونكتب:

$$T \rightsquigarrow t_v$$

[1]-جنيدي مراد، "**محاضرات في الإحصاء 3**"، مرجع سبق ذكره، ص38.

1-2-المميزات العددية للتوزيع:

ان المميزات العددية للتوزيع ستيودنت تتمثل في:[1]

-التوقع الرياضي:

$$E(T)=0$$

-التباين:

$$V(T)=\frac{v}{v-2} \quad si\ v>2$$

1-3-التمثيل البياني للتوزيع:

ان التمثيل البياني لتوزيع ستيودنت يتميز بـ:

-هو توزيع احتمالي يشبه في شكله نوعا ما التوزيع الطبيعي، الا انه يتميز بذيول سميكة ما يجعله ملائما للعينات الصغيرة؛

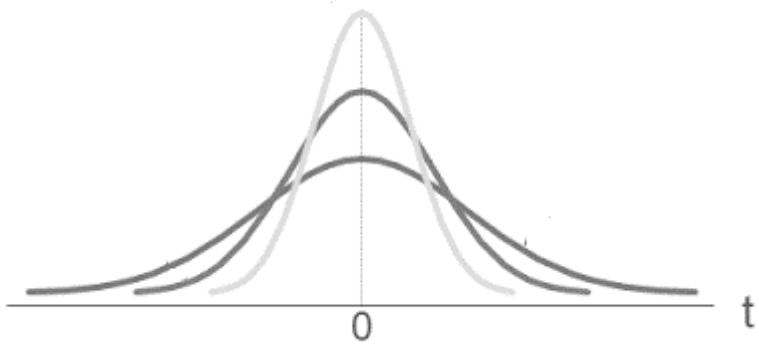
-شكله يأخذ شكل الجرس المتماثل عند الصفر (0)، هذا يعني ان كل من المتوسط والوسيط والمنوال كلها متساوية وتساوي الصفر؛

-يعتمد في شكله على درجات الحرية "*Degrees of Freedom - df*" بشكل كبير، والتي ترتبط بدورها بحجم العينة "*df=n-1*"؛

-ان كبر درجات الحرية تقرب التمثيل البياني لتوزيع ستيودنت من التمثيل البياني للتوزيع الطبيعي والقياسي "*Standard Normal Distribution*"؛

-عندما تنخفض درجة الحرية فان التمثيل البياني لتوزيع ستيودنت يكون مسطحا أكثر في منتصفه ومرتفعا عند الذيول، هذا ما يدل على وجود قيم شاذة او متطرفة.

[1]-نفس المرجع السابق، ص38.



2-التوزيع فيشر:

يعد توزيع فيشر من التوزيعات الإحصائية الهامة، حيث يستخدم في الإحصاء الاستنتاجي لإجراء العديد من اختبارات الفروض المتعلقة بتحليل التباين، وتصميم التجارب واختبار معنوية خطوط الانحدار، الى غير ذلك من التطبيقات الإحصائية العديدة والهامة، ولقد اكتشف هذا التوزيع من قبل العالم الإنجليزي الشهير "رونالد آيلمر فيشر" "*Ronald Aylmer Fisher*"، حيث استخدمه لاختبار النسبة بين تبايني مجتمعين طبيعيين.[1]

ان الاستخدامات المهمة لتوزيع فيشر هي:

-يستخدم لتحليل التباين، وهو الاستخدام الأكثر شيوعا، لتحديد هل هناك فروق معنوية بين متوسطات عدة مجموعات، ويعتمد عليه لمقارنة التباينات بين المجموعات، وأيضا التباين داخل المجموعات؛

-ويستخدم في اختبار المعنوية الكلية لنماذج الاقتصاد القياسي، من اجل قبولها او رفضها كليا، وتكون ذات معنوية إحصائية، وقابلة للتحليل الاقتصادي فيما بعد؛

-ويستخدم من اجل اجراء اختبار تساوي تباين مجتمعين مستقلين "*F-Test for Equality of Variances*"، لأنه يعتبر مهم قبل اجراء أي اختبار اخر على المجتمعات؛

- يستخدم في "اختبار الميزات" "*Feature Selection*"، الذي بواسطته يتم اختبار وتحديد المتغيرات الأكثر تأثير في دقة التوقعات لتقليل تعقيد النماذج المبرمجة، في علم البيانات والذكاء الاصطناعي.

[1]-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، مرجع سبق ذكره، ص393.

2-1-التعريف:[1]

تعريف توزيع فيشر:

هو أحد التوزيعات الإحصائية المتصلة المهمة والمستخدمة في اختبار الفرضيات وفي تحليل التباين يعرف كما يلي:

إذا كان "X, Y" متغيرين عشوائيين،وكان "X" له توزيع كاي مربع بدرجات حرية "v_1"، والمتغير العشوائي "Y" له توزيع كاي مربع بدرجات حرية "v_2"، فان المتغير العشوائي:

$$F=\frac{\frac{X}{v_2}}{\frac{Y}{v_1}}$$

ان المتغير العشوائي "F" يتبع توزيع فيشر بدرجات حرية "v_1, v_2".

ان دالة كثافة المتغيرة العشوائية "F" تكتب على الشكل الموالي:

$$f(F)=\frac{C.F^{\frac{v_1-2}{2}}}{(v_2+v_1F)^{\frac{v_1+2}{2}}} \qquad F>0$$

حيث ان "C" هو ثابت يعتمد على "v_1, v_2"، ليجعل المساحة تحت المنحنى تساوي الواحد (1).

ونقول ان المتغيرة "F" تتبع التوزيع فيشر بـ "v_1, v_2"درجة حرية ونكتب:

$$F \rightsquigarrow f_{v_1, v_2}$$

[1]-بوزنورة اسماء، "**الإحصاء (02) -دروس وتمارين-**"، مرجع سبق ذكره، ص91.

1-2-المميزات العددية للتوزيع:

ان المميزات العددية لتوزيع فيشر تتمثل في:

-التوقع الرياضي:

$$E(F)=\frac{v_2}{v_2-2} \quad v_2>2$$

-التباين:

$$V(F)=\frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-4)(v_2-2)^2} \qquad v_2>4$$

1-3-التمثيل البياني للتوزيع:

ان التمثيل البياني لتوزيع فيشر يستعمل بشكل أساسي في اختبار الفرضيات، ويعتمد الشكل البياني له على درجات الحرية لكل من البسط والمقام، وهو يتميز بـ:

-التمثيل البياني لتوزيع فيشر هو عبارة عن منحنى ليس متماثل، ويكون ملتويا نحو اليمين "*Positively Skewed*"؛

-ان التمثيل البياني لتوزيع فيشر لا يعتمد على القيم السالبة، ويبدأ من القيمة الصفر (0)، ويمتد الى ما لا نهاية "$+\infty$"؛

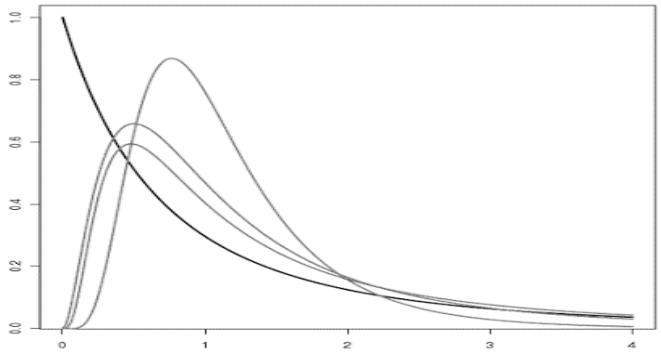
-يعتمد ويتأثر التمثيل البياني لتوزيع فيشر بشكل كبير على درجات الحرية كما يلي:

-كلما كانت درجات الحرية صغيرة كلما كان التمثيل البياني لتوزيع فيش أكثر تشتتا والتواء؛

-وكلما زادت او كبرت درجات الحرية كلما اقترب المنحنى من منحنى التوزيع الطبيعي ويقترب الى التماثل تدريجيا.

-ان قمة توزيع فيشر تكون قريبة من القيمة صفر كلما كانت درجات الحرية صغيرة؛

-ان التمثيل البياني لتوزيع فيشر يستخدم لقبول او رفض الفرضيات في اختبارات فيشر.



3-تمارين محلولة:[1]

التمرين الأول

أ-ما هي قيمة المساحة التي تقابل t_{11}=1.812؛

ب-اوجد قيمة t_7 على يمينها مساحة 0.01؛

ج-اوجد قيمة t_3 على يسارها مساحة قدرها 0.95.

أ-قيمة المساحة التي تقابل t_{11}=1.812:

لدينا:

$v=n-1$

$n=11$

$v=11-1$

$v=10$

باستعمال جدول توزيع ستيودنت فان المساحة المقابلة هي 0.05.

ب-ايجاد قيمة t_7 على يمينها مساحة 0.01:

$P(t_7>t)=0.01$

من جدول ستودنت نجد ان:

$X=2.998$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

ج-اوجد قيمة t_3 على يسارها مساحة قدرها 0.95:

$P(t_3 < x)=0.95$

$P(t_3 > t)=1-0.95$

$P(t_3 > t)=0.05$

هنا انتقلنا من الجانب الايسر الى الجانب الأيمن، وبالاستعانة بجدول ستيودنت نجد القيمة:

$X=2.353$

ونكتب:

$P(t_3 > 2.353)=0.05$

او:

$P(t_3 < 2.353)=0.95$

التمرين الثاني

ان متوسط انتاج شركة صناعية للمواد البلاستيكية في الأسبوع من الاطقم البلاستيكية هو "$\mu=825$" طقما وان التباين معلوم، وللتأكد من هذه المعطيات، اختيرت عينة عشوائية ذات الحجم "$n=25$" من منتجات هذه الشركة، فكان متوسط الإنتاج الأسبوعي هو "$\overline{X}=780$" طقما وبانحراف معياري "$\sigma=105$".

فاذا علمت ان منتجات الشركة تتوزع توزيعا طبيعيا على وجه التقريب اوجد احتمال ان يكون متوسط الإنتاج الأسبوعي للشركة 780 على الأكثر.

$$P(\overline{X}\leq 780)=P(t_{24}\leq \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}})$$

$$P(\overline{X}\leq 780)=P(t_{24}\leq \frac{780-825}{\frac{105}{\sqrt{25}}})$$

$$P(\overline{X}\leq 780)=P(t_{24}\leq -2.143)$$

$$P(\overline{X}\leq 780)=1-P(t_{24}\geq 2.143)$$

بما ان الدرجة المعيارية في جدول توزيع ستيودنت 2.143 لا تظهر لذا يمكننا تقديرها بالطريقة التالية:

$P(t_{24} \leq 2.064) = 0.975$

$P(t_{24} \leq 2.492) = 0.990$

هذا يعني ان:

$P(t_{24} \leq 2.064 \leq P(t_{24} \geq 2.143) \leq P(t_{24} \leq 2.492)$

$0.01 \leq P(\bar{X} \leq 780) \leq 0.025$

أي ان الحصول على 780 طقما في الأسبوع على الأكثر يحدث بنسبة 1% الى 2.5%.

التمرين الثالث

قم بإيجاد ما يلي:

أ- $f(v_1, v_2) = f_{(5,7)(0.01)}$؛

ب- $f_{(9,10)(0.01)}$.

أ-إيجاد $f(v_1, v_2) = f_{(5,7)(0.01)}$:

من الجدول الاحتمالي لتوزيع فيشر فان:

$f_{(5,7)(0.01)} = 7.49$

ب-إيجاد $f_{(9,10)(0.01)}$:

من الجدول الاحتمالي لتوزيع فيشر فان:

$f_{(9,10)(0.01)} = 4.94$

التمرين الرابع

في تجربة لدراسة مدخول الافراد ضمن مجتمعين احصائيين طبيعيي التوزيع، تم اخذ عينتين ذواتي حجمين "$n_1=11,\ n_2=16$" وكان تباين كل من العينتين "$\sigma_1^2=(2000)^2,\ \sigma_2^2=(2400)^2$"، احسب كلا من:

- $F_{0.95,\ 10,\ 15}$؛
- $F_{0.99,\ 10,\ 15}$؛
- $F_{0.05,\ 10,\ 15}$؛
- $F_{0.01,\ 10,\ 15}$؛
- $F_{0.95,\ 15,\ 10}$؛
- $F_{0.99,\ 15,\ 10}$؛
- $F_{0.05,\ 15,\ 10}$؛
- $F_{0.01,\ 15,\ 10}$؛

من جداول توزيع فيشر فان:

$F_{0.05,\ 10,\ 15}=2.54$

$F_{0.01,\ 10,\ 15}=3.80$

ومنه:

$$F_c=\frac{\sigma_1^2}{\sigma_2^2}$$

$$F_c=\frac{(2000)^2}{(2400)^2}$$

$$F_c=0.69$$

$$P(F_c\leq F=3.80)=1-\alpha$$

$$P(F_c\leq F=3.80)=0.99$$

$$P(F_c\leq F=2.54)=1-\alpha$$

$$P(F_c\leq F=2.54)=0.95$$

$$F_{0.99,\ 15,\ 10}=\frac{1}{F_{0.01,\ 10,\ 15}}$$

$$F_{0.99,\ 15,\ 10}=\frac{1}{4.56}$$

$$F_{0.99,\ 15,\ 10}=0.22$$

$$F_{0.95,\,10,\,15} = \frac{1}{F_{0.05,\,15,\,10}}$$

$$F_{0.95,\,10,\,15} = \frac{1}{2.85}$$

$$F_{0.95,\,10,\,15} = 0.35$$

$$F_{0.99,\,15,\,10} = \frac{1}{3.80}$$

$$F_{0.95,\,15,\,10} = \frac{1}{2.54}$$

$$P\left(F_{0.99,10,15} = 0.22 < F_c = 0.69 < 3.80 = F_{0.01,10,15}\right) \geq 0.98 = 1-2\alpha$$

$$P\left(F_{0.95,10,15} = 0.35 < F_c = 0.69 < 2.54 = F_{0.05,10,15}\right) \geq 0.90 = 1-2\alpha$$

المحاضرة الثانية عشر

تقارب بعض التوزيعات الاحتمالية

المحاضرة الثانية عشر: تقارب بعض التوزيعات الاحتمالية

ان:

-تقارب التوزيعات الاحتمالية "*Convergence of probability distributions*": يرجع ظهورها الى القرن الثامن عشر مع أعمال عالم الرياضيات الفرنسي "أبراهام دي موافر" "*Abraham de Moivre*" (1667-1754)، الذي يعتبر هو اول من استخدم التقارب بين التوزيعات الاحتمالية، من خلال تقريب توزيع ذي الحدين الى التوزيع الطبيعي، ليأتي بعده عالم الرياضيات والاحصاء الفرنسي "بيير سيمون ماركيز دو لابلاس" "*Pierre Simon marquis de Laplace*" (1749-1827) الذي وضع الأساس وأرسى مبادئ "نظرية النهاية المركزية" "*Central Limit Theorem*"، والتي تعتبر حجر الزاوية لتقارب عدد كبير من التوزيعات الاحتمالية. ليأتي فيما بعد مجموعة كبيرة من العلماء الذين قاموا بتطوير وتوسيع نطاق التقارب بين التوزيعات الاحتمالية، اهمهم عالم الرياضيات الروسي "بافنوتي شبيشيف" "*Pafnuty Chebyshev*" (1821-1894).

1-تقارب التوزيعات الاحتمالية "*Probability Distribution Approximation*":

ان التوزيعات الاحتمالية لها استخدامات مختلفة في عملية اتخاذ القرار، التي هي قائمة على بيانات عملية، ونستطيع ان نعدد بعض الاستخدامات العملية للتوزيعات الاحتمالية فيما يلي:

-توزيع بواسون "*Poisson Distribution*" يستخدم بشكل كبير في تقدير الحوادث او تحديد الأقساط التأمينية التي يطالب بها المؤمنون على ممتلكاتهم من شركات التامين، وهو يساعد شركات التامين في التدقيق في سياسات الإدارة للمخاطر التي تواجهها مستقبلا؛

-يعتمد مهندسو الإنتاج والتصنيع على التوزيع الطبيعي "*Normal Distribution*" من اجل تحديد العيوب في الوحدات المنتجة وخطوطها الإنتاجية، وبالتالي القيام بعمليات الإصلاح اللازمة؛

-تسعى شركات الاتصال الى تصميم وتحسين شبكات الاتصال من اجل تحمل الضغط عليها، ولأجل ذلك فإنها تستعين بالتوزيع الأسي "*Exponential Distribution*"؛

-ولأجل نمذجة نتائج العلاج على مستوى المؤسسات الصحية للتجارب السريرية، فإنها تستخدم توزيع برنولي "*Bernoulli Distribution*".

مما سبق فان استعمالات التوزيعات تتعدد بتعدد المجالات التي تستخدمها، هذا ما يجعل من هذه التوزيعات تتعقد، ويصعب على مستعمليها تحديد الاحتمالات المتعلقة بها، من هنا نستطيع ان نعطي تعريف بسيط لتقارب التوزيعات الاحتمالية "*Probability Distribution Approximation*"، كما يلي:

هو "**عملية استعمال او الاستعانة بتوزيع احتمالي بسيط يحل محل توزيع اخر أكثر تعقيدا منه، من اجل تسهيل حساب القياسات الإحصائية والرياضية**".

يتم الاستعانة بتقارب التوزيعات الاحتمالية لأجل تبسيط حل المشاكل الصعبة باستخدام توزيع ابسط من التوزيع الأصلي الذي يتناسب معها، ومع توفر شرط كبر حجم العينة.

2-أنواع التقارب الاحتمالي:

ان أنواع التقارب الاحتمالي الرئيسية هي: [1]

2-1-التقارب في التوزيع "*Convergence in Distribution*":

إذا كانت "X_n" سلسلة من المتغيرات العشوائية معرفة على فضاء احتمالي "Ω"، حيث ان "F_n" تمثل دوال التوزيع الخاصة بها، كما ان "X" متغير عشوائي معرف على نفس الفضاء الاحتمالي دالة توزيعه "F".

نقول ان سلسلة المتغيرات العشوائية "X_n" تتقارب في القانون الى "X" إذا وفقط كل نقطة "x" تكون عندها دالة التوزيع "F" مستمرة، اين:

$$\lim_{n\to\infty} F_n(x)=F(x)$$

وحسب "نظرية موافر-لابلاس" "*Moivre-Laplace theorem*"، نعتبر ان "X_n" سلسلة من المتغيرات العشوائية التي تتبع التوزيع الثنائي "$B(n,p)$"، حيث ان "p" معلمة ثابتة، يمكن ان نقول:

$$Z=\frac{X_n-np}{\sqrt{npq}}$$

يتقارب الى القانون الطبيعي المعياري.

2-2-التقارب في الاحتمال "*Convergence in Probability*":

إذا كانت "X_n" سلسلة من المتغيرات العشوائية معرفة على فضاء احتمالي "Ω"، كما ان "X" متغير عشوائي معرف على نفس الفضاء الاحتمالي.

نقول ان سلسلة المتغيرات العشوائية "X_n" تتقارب في الاحتمال الى "X"، إذا وفقط مهما يكن "ε" عدد حقيق موجب:

$$\lim_{n\to\infty} P(|X_n-X|<\varepsilon)=1$$

[1]-عبدلي ادريس، "**دروس عبر الخط في مقياس الإحصاء 03**"، مرجع سبق ذكره، ص68، 70.

وحسب متراجحة "بافنوتي شبيشيف" "*Pafnuty Chebyshev*" فاذا كان "X" متغير عشوائي، بمتوسط "μ"، وانحراف معياري "σ"، فمن اجل كل عدد حقيقي موجب "k" فان:

$$P(|X-\mu|\geq k\sigma)\leq\frac{1}{k^2}$$

3-نظرية النهاية المركزية:[1]

لتكن المتغيرات "$X_1,X_2,\ldots,X_n$" متغيرات عشوائية مستقلة، لها نفس التوزيع الاحتمالي بتباين ومتوسط محددين، إذا كانت:

$$S_n=X_1+X_2+\ldots+X_n$$
$$n=1,2,\ldots$$

فان:

"S_n" تتبع التوزيع الطبيعي عندما:

$$n\longrightarrow\infty$$

وبما ان:

$$E(S_n)=n\mu$$
$$\sigma_{S_n}=\sigma\sqrt{n}$$

فإننا نكتب:

$$\lim_{n\to\infty}P\left(a\leq\frac{S_n-n\mu}{\sigma\sqrt{n}}\leq b\right)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{\frac{-Z^2}{2}}dZ$$

ان الحقيقة النظرية محققة عندما تكون المتغيرات العشوائية مستقلة "$X_1,X_2,\ldots,X_n$"، ولها نفس المتوسط والتباين حتى ولو لم يكن لها بالضرورة نفس التوزيع، مع العلم انه توجد صيغ أخرى لهذه النظرية، حيث لا يشترط ان يكون للمتغيرات نفس التوزيع الاحتمالي، ولا حتى ان تكون مستقلة.

[1]-جنيدي مراد، "**محاضرات في الإحصاء 3**"، مرجع سبق ذكره، ص43، 44.

وتعتبر نظرية "موافر-لابلاس" حالة خاصة من نظرية النهاية المركزية، ذلك ان المتغيرة العشوائية تتبع قانون "$B(n,p)$"، ويمكن اعتبارها مجموعا لعدد من المتغيرات العشوائية المستقلة ذات التوزيع برنولي "$B(1,p)$ ".

4-استمرارية تقارب التوزيعات الاحتمالية:

ان استمرارية تقارب التوزيعات الاحتمالية، هي عبارة عن امكانية تقريب او الانتقال من توزيع احتمالي منفصل الى توزيع اخر متصل، والعكس صحيح، مع توفر ظروف كبر حجم العينة، لكي يتسنى لمستخدمي التوزيعات الاحتمالية من حساب بعض الاحتمالات المعقدة.

ان معالجة حالة استمرارية تقارب التوزيعات الاحتمالية تتم وفق ما يلي:[1]

عند تقريب توزيع احتمالي منفصل الى توزيع احتمالي متصل فانه ينبغي القيام بتصحيح الاستمرارية، وذلك لان هستم الانتقال من متغيرة عشوائية منفصلة الى متغيرة عشوائية متصلة، لذلك يتم تحويل كل قيمة للمتغير العشوائي الى مجال، لان في التوزيع المستمر يكون احتمال قيمة معينة الصفر (0)، وتتم عملية التحويل بحيث نزيل التقطع والانفصال بين القيم (اعداد طبيعية)، ويكون ذلك بإضافة 0.5، وطرح 0.5 من كل قيمة من قيم المتغير العشوائي المنفصل، فنفرض مثلا:

لدينا قيم المتغير العشوائي المنفصل التالي:

$\{1, 2, 3\}$

فتصبح مجالات المتغير العشوائي المستمر هي:

$\{0.5\text{-}1.5, 1.5\text{-}2.5, 2.5\text{-}3.5\}$

وبالتالي نلاحظ انه أصبحت هناك استمرارية، أي انه في الحالة العامة عندما يكون المتغير العشوائي يساوي قيمة "a"، فانه يتم استبدال هذه القيمة بالمجال:

$[(a\text{-}0.5);(a+0.5)]$

اذن ففي التوزيعات التي تم دراستها ينبغي القيام بتصحيح الاستمرارية، والتي تتضمن إضافة او طرح 0.5، على حسب الحالة، ففي الحالة العامة يكون ذلك كالاتي:

$\mathbf{-P(X=a)}=P(a\text{-}0.5\leq X\leq a+0.5)$

[1]- حمبلي زهير، "**دروس عبر الخط في مقياس الإحصاء 03**"، جامعة ام البواقي، ص105، 106.

-P(X<*a*)=P(X<**[**(*a*-0.5);(*a*+0.5)**]**)

P(X<*a*)=P(X<*a*-0.5)

-P(X≤*a*)=P(X≤**[**(*a*-0.5);(*a*+0.5)**]**)

P(X<*a*)=P(X≤*a*+0.5)

-P(X>*a*)=P(X>**[**(*a*-0.5);(*a*+0.5)**]**)

P(X>*a*)=P(X>*a*+0.5)

-P(X≥*a*)=P(X≥**[**(*a*-0.5);(*a*+0.5)**]**)

P(X≥*a*)=P(X≥*a*-0.5)

-P(*a*<X<*b*)=P(**[**(*a*-0.5);(*a*+0.5)**]**<X<**[**(*b*-0.5);(*b*+0.5)**]**)

P(*a*<X<*b*)=P(*a*+0.5<X<*b*-0.5)

-P(*a*≤X≤*b*)=P(**[**(*a*-0.5);(*a*+0.5)**]**≤X≤**[**(*b*-0.5);(*b*+0.5)**]**)

P(*a*≤X≤*b*)=P(*a*-0.5≤X≤*b*+0.5)

-P(*a*<X≤*b*)=P(**[**(*a*-0.5);(*a*+0.5)**]**<X≤**[**(*b*-0.5);(*b*+0.5)**]**)

P(*a*<X≤*b*)=P(*a*+0.5<X≤*b*+0.5)

-P(*a*≤X<*b*)=P(**[**(*a*-0.5);(*a*+0.5)**]**≤X<**[**(*b*-0.5);(*b*+0.5)**]**)

P(*a*≤X<*b*)=P(*a*-0.5≤X<*b*-0.5)

ان في حالة عندنا تجربة تتبع توزيع متصل في الأصل، او تتبع توزيع منفصل وقمنا بتقريبها لتوزيع متصل، فانه يكون عندنا احتمال مجال فقط، ولا يكون هندنا احتمال قيمة معينة، لان الاحتمال عند قيمة معينة يساوي الصفر (0)، أي:

$$P(X=a)=0$$

وبالتالي ينتج عن هذا في الحالة العامة عند التوزيع المتصل يكون:

$$P(a\leq X\leq b)=P(a<X\leq b)=P(a\leq X<b)=P(a<X<b)$$

أي انه في حالة التوزيعات المتصلة (مثل التوزيع الطبيعي)، فانه عند حساب الاحتمال في مجال معين، فانه لا يهم ان يكون مغلق او مفتوح من أحد الطرفين او كلاهما، لان النتيجة ستكون نفسها.

5-بعض أشهر تقارب التوزيعات الاحتمالية:

يوجد عدد كبير من التقريبات بين التوزيعات الاحتمالية، سوف نذكر مجموعة منها:[1]

5-1-تقارب التوزيع ذو الحدين بالتوزيع الطبيعي:

إذا كان "X" متغير عشوائي يمثل عدد مرات نجاح الحدث، عند اجراء تجربة برنولي لمرات متعددة تساوي "n" مرة، فانه يأخذ القيم:

$$X=\{0, 1, 2, \ldots, n\}$$

باحتمال قدره:

$$P(X=x)$$

يحسب كالتالي:

$$f(x)=P(X=x)=\begin{cases} C_n^x P^x (1-p)^{n-x} & x=0, 1, 2, \ldots, n \\ 0 & or \end{cases}$$

$$P(X=x)=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

حيث ان:

-"n": هو عدد مرات تكرار التجربة؛

-"$P(X=x)$": احتمال ان يكون المتغير العشوائي "X" يأخذ القيمة "x"؛

-"p": احتمال النجاح.

تقارب التوزيع ذو الحدين بالتوزيع الطبيعي هو عملية تقدير الاحتمالات المرتبطة بتوزيع ذي الحدين، عندما "$n\longrightarrow\infty$" فان توزيع المتغير العشوائي:

$$Z=\frac{x-np}{\sqrt{np(1-p)}}$$

[1]-مزواغي جيلالي، "**إحصاء 3**"، مطبوعة بيداغوجية، جامعة غيليزان، 2025/2024، ص88، 94.

هذه النظرية اثبتها أولا الرياضي الفرنسي "أ.ديموفري" (1733)، يتيح لنا استخدام هذه النظرية عندما يكون "n" كبيرا، تقليل مجموعات الاحتمالات المرتبطة بتوزيع ذي الحدين من خلال تكاملات مناسبة للتوزيع الطبيعي القياسي. بشكل أكثر دقة، تنص النظرية على انه عندما يكون "n" كبيرا، فان المتغير العشوائي "X"يتبع تقريبا التوزيع الطبيعي "$N(np,np(1-p))$"، يتحسن التقريب كلما زاد "n"، وهو جيد جدا للقيم التي لا تقترب كثيرا من الصفر (0) او الواحد (1).

علاوة على ذلك، يكون التقريب أفضل عندما يكون:

$$\begin{cases} np>5 \\ (1-p)>5 \end{cases}$$

5-2-تقارب توزيع بواسون بالتوزيع الطبيعي:

إذا كان "X" متغيرا عشوائيا يتبع توزيع بواسون بمتوسط وتباين "λ":

$$P(a\leq x\leq b)=P(\frac{(a-0.5)-\lambda}{\sqrt{\lambda}}\leq Z\leq\frac{(a+0.5)-\lambda}{\sqrt{\lambda}})$$

حيث ان:

-"a, b": هما عددان صحيحان غير سالبين؛

-"Z": هو متغير عشوائي طبيعي قياسي؛

-"$0.5\mp$": هو يستخدم معامل تصحيح الاستمرارية.

يستخدم معامل تصحيح الاستمرارية "$0.5\mp$" لتحسين دقة التقريب عند الانتقال من التوزيع المنفصل الى التوزيع المتصل.

5-3-تقارب توزيع ذو الحدين بتوزيع بواسون:

هو تقنية شائعة تستخدم عندما تكون عدد التجارب في توزيع ذي الحدين كبيرا، ونجاح التجربة "*Probability of success*" صغيرا.

لنفترض ان لدينا توزيع ذي الحدين "X" حيث "n" عدد التجارب (معدل كبير)، و"p" احتمال النجاح في كل تجربة (معدل صغير).

يمكن تقارب بواسون عندما يكون "n" كبيرا، و"p" صغيرا، يمكننا استخدام توزيع بواسون كما يلي:

$$P(X=k)=\frac{e^{-\lambda}\lambda^{k}}{k!}$$

5-4-تقارب توزيع كاي مربع بالتوزيع الطبيعي:

عندما تكون درجات الحرية "n" كبيرة، يمكن استخدام التقريب التالي:

يمكن اعتبار توزيع كاي مربع كجمع لمربعات "k" متغيرات طبيعية مستقلة، مما يعني انه يمكن تقريبه بالتوزيع الطبيعي:

$$\chi_n^2 \rightsquigarrow N(n,\ 2n)$$

6-تمارين محلولة:[1]

التمرين الأول

ليكن "X" متغيرا عشوائيا يتبع توزيع ثنائي الحدين مع الدالة الاحتمالية التالية:

$$f(x)=P(X=x)=\begin{cases} C_{25}^{x}\left(\frac{1}{4}\right)^{x}\left(\frac{3}{4}\right)^{25-x} & x=0,\ 1,\ 2,\ \ldots,\ 25 \\ 0 & or \end{cases}$$

المطلوب حساب "$P(2\leq x\leq 8)$".

لدينا:

$$n=25$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

$$p=\frac{1}{4}$$

$$\mu=np$$

$$\mu=25\times\frac{1}{4}$$

$$\mu=6.25$$

$$\sigma^2=np(1-p)$$

$$\sigma^2=25\times\frac{1}{4}\times(1-\frac{1}{4})$$

$$\sigma^2=4.68$$

التقريب باستخدام التوزيع الطبيعي:

$$X\rightsquigarrow N(np,np(1-p))$$

$$X\rightsquigarrow N(6.25,4.68)$$

$$f(x)=P(2\leq x\leq 8)$$

$$P(2\leq x\leq 8)=\sum_{x=2}^{8} C_{25}^{x}(\frac{1}{4})^{x}(\frac{3}{4})^{25-x}$$

$$Z=\frac{x-np}{\sqrt{np(1-p)}}$$

$$P(2\leq x\leq 8)=P(\frac{x-np}{\sqrt{np(1-p)}}\leq Z\leq\frac{x-np}{\sqrt{np(1-p)}})$$

$$P(2\leq x\leq 8)=P(\frac{2-6.25}{4.33}\leq Z\leq\frac{8-6.25}{4.33})$$

$$P(2\leq x\leq 8)=P(-0.98\leq Z\leq 0.40)$$

$$P(2\leq x\leq 8)=0.4919$$

التمرين الثاني

نفرض انه عندنا تجربة احتمالية مكررة تتمثل في رمي قطعة نقود 33 مرة، حيث احتمال ان يظهر وجه في كل تجربة هو 0.6، واحتمال ظهور الظهر هو 0.4، وان "X" هو متغير عشوائي يمثل عدد الأوجه التي يمكن ان تظهر، ونريد حساب احتمال ان يظهر الوجه أكثر من 14 مرة.

نلاحظ ان هذه التجربة الاحتمالية المكررة تتبع قانون ثنائي الحد، لكن عند استخدامه فان القيام بالعمليات الحسابية تكون صعبة جدا:

$$P(X>14)=P(15)+P(16)+...+P(33)$$

$$P(X>14)=C_{33}^{14}\times(0.6)^{14}\times(0.4)^{19}+...$$

نلاحظ انه يصعب القيام بالعمليات الحسابية لحساب هذا الاحتمال، لذلك يتم استخدام قانون التوزيع الطبيعي كتقريب لقانون ثنائي الحد، لان شروط التقريب متوفرة، حيث:

$$n=33\geq 30$$

$$np=33\times 0.4$$

$$np=13.2$$

$$np\geq 5$$

$$n(1-p)=33\times 0.6$$

$$n(1-p)=19.8\geq 5$$

وعند استخدام التوزيع الطبيعي، بحيث:

$$\sigma^2=np(1-p)$$

$$\sigma^2=33\times 0.6\times 0.4$$

$$\sigma^2=7.92$$

$$\mu=np$$

$$\mu=33\times 0.6$$

$$\mu=19.8$$

بتصحيح الاستمرارية، يصبح الاحتمال المطلوب هو:

$P(X>14.5)$

ويتم حسابه كالاتي:

$$P(X>14.5)=P(Z>\frac{14.5-19.8}{\sqrt{7.92}})$$

$$P(X>14.5)=P(Z>-1.88)$$

$$P(X>14.5)=P(-1.88<Z<0)+P(0<Z<+\infty)$$

$$P(X>14.5)=0.4699+0.5$$

$$P(X>14.5)=0.9699$$

التمرين الثالث

نفترض ان عدد المكالمات الواردة الى مركز خدمة العملاء في الساعة يتبع توزيع بواسون، بمتوسط 10، نريد حساب احتمال ان يتلقى المركز بين 8 و12 مكالمة في الساعة الواحدة.

حسب قانون بواسون:

$$P(8\leq X\leq 12)=\sum_{x=8}^{12}\frac{e^{-\lambda}\lambda^{X}}{X!}$$

$e=2.718$

متوسط عدد المكالمات الواردة الى مركز خدمة العملاء في الساعة هو:

$\lambda=10$

$$P(8\leq X\leq 12)=\sum_{x=8}^{12}\frac{e^{-\lambda}\lambda^{X}}{X!}$$

$$P(8\leq X\leq 12)=0.1120+0.1240+0.1240+0.1136+0.0947$$

$$P(8\leq X\leq 12)=0.8683$$

التقريب باستخدام التوزيع الطبيعي:

$X\rightsquigarrow N(\lambda, \sqrt{\lambda})$

$X \rightsquigarrow N(10, \sqrt{10})$

$$P(8 \leq X \leq 12) = P\left(\frac{(a-0.5)-\lambda}{\sqrt{\lambda}} \leq Z \leq \frac{(a+0.5)-\lambda}{\sqrt{\lambda}}\right)$$

$$P(8 \leq X \leq 12) = P\left(\frac{(8-0.5)-10}{3.1623} \leq Z \leq \frac{(12+0.5)-10}{3.1623}\right)$$

$$P(8 \leq X \leq 12) = P(-0.79 \leq Z \leq 0.79)$$

$$P(8 \leq X \leq 12) = 0.5704$$

التمرين الرابع

وكالة سياحية تستقبل في المتوسط 12.5 مكالمة هاتفية كل 5 دقائق، فما احتمال ان تستقبل بين 10:00 و 10:05 كل من:

-اقل من 9 مكالمات هاتفية؛

-9 مكالمات هاتفية الأكثر؛

-10 مكالمات؛

-أكثر من 15 مكالمة هاتفية؛

-أكثر من 5 واقل من او يساوي 15 مكالمة هاتفية.

نلاحظ ان هذه التجربة الاحتمالية تتبع في الأصل توزيع بواسون، حيث:

$\lambda = 12.5$

حيث يكون:

$\mu = \sigma^2 = \lambda = 12.5$

$-P(X<9) = P(X<8.5)$

$$P(X<9) = P\left(Z < \frac{8.5-12.5}{\sqrt{12.5}}\right)$$

$P(X<9) = P(Z<-1.13)$

$P(X<9) = P(Z>1.13)$

$P(X<9) = 0.5 - P(0<Z<1.13)$

$P(X<9) = 0.5 - 0.3708$

P(X<9)=0.1292

-P(X≤9)=P(X<9.5)

$$P(X\leq 9)=P(Z<\frac{9.5-12.5}{\sqrt{12.5}})$$

P(X<9)=0.1977

-P(X=10)=P(9.5<X<10.5)

$$P(X=10)=P(\frac{9.5-12.5}{\sqrt{12.5}}<Z<\frac{10.5-12.5}{\sqrt{12.5}})$$

P(X<9)=0.09

-P(X>15)=P(X>15.5)

$$P(X>15)=P(Z>\frac{15.5-12.5}{\sqrt{12.5}})$$

P(X<9)=0.1977

-P(5<X≤15)=P(5.5<X<15.5)

$$P(5<X\leq 15)=P(\frac{4.5-12.5}{\sqrt{12.5}}<Z<\frac{15.5-12.5}{\sqrt{12.5}})$$

P(5<X≤15)=0.7904

المحاضرة الثالثة عشر

المتغيرات العشوائية المزدوجة

المحاضرة الثالثة عشر: المتغيرات العشوائية المزدوجة

ان:

– المتغيرات العشوائية المزدوجة "*Bivariate random variables*": ليس لها تاريخ او فترة محددة لظهورها، ولا يوجد عالم او علماء محددين الذين ساهموا في تأسيسها، بل تطور هذا المصطلح بمفهومه تدريجيا مع تطور علم او نظرية الاحتمالات بين القرنين الثامن عشر (17)، والقرن التاسع عشر (18)، لنمذجة الظواهر التي تعتمد على أكثر من حدث او متغير، وتعتبر من الأسس الهامة في الاقتصاد القياسي لبناء العلاقات بين المتغيرات. فلقد نشا هذا المصطلح مع دراسة العاب الحظ من طرف كل من عالم الفيزياء والرياضيات الفرنسي "بليز باسكال" "*Blaise Pascal*" (1623-1662) وعالم الرياضيات الفرنسي "بيير فارمات" "*Pierre de Fermat*" (1601-1665)، لياتي بعدهما "جاكوب برنولي" "*Jakob Bernoulli*" (1654-1705) ليضع أسس الاحتمالات، هذا ما مهد لظهور دراسات تهتم بالأحداث او الظواهر التي لها اكثر من نتيجة، ومع ظهور اعمال كل من عالم الرياضيات الفرنسي "ابراهام دي موافر" "*Abraham de Moivre*" (1667-1754) وعالم الرياضيات والاحصاء الألماني "يوهان كارل فريديريك غوس" "*Johann Carl Friedrich Gauss*" (1777-1855)، اللذان ركزا وطورا استعمال التوزيع الطبيعي، وبالخصوص في تحديد القياسات المتعددة، وبين القرن العشرين (19) واوائل القرن الواحد والعشرين (20)، عرف المصطلح تطورا كبيرا من خلال اعمال كل من عالم الرياضيات الروسي "اندريه ماركوف" "*Andrey Andreyevich Markov*" (1856-1922)، وعالم الرياضيات الروسي "بافنوتي شبيشيف" "*Pafnuty Chebyshev*" (1821-1894)، اللذان بدورهما اسسا الى ظهور مفاهيم جد مهمة في علم القياس الاقتصادي، التي تتعلق بالارتباط "*Correlation*"، والانحدار "*Regression*"، التي تعتمد على العلاقة في ان واحد بين متغيرتين على الأقل في نمذجتها، وبعدها عرف المصطلح تطور كبير بتطور أدوات الاقتصاد القياسي.

1-المتغيرات العشوائية المزدوجة:

المتغيرات العشوائية الثنائية او المزدوجة، تشير الى مجموعة من المتغيرات العشوائية التي يتم دراستها معا لتحديد العلاقة او الارتباط بينها. المتغير العشوائي الثنائي هو نوع من المتغيرات العشوائية التي تتعامل مع حالتين او أكثر في نفس الوقت، حيث يتم دراسة العلاقة بين متغيرتين عشوائيتين. يستخدم هذا النوع من المتغيرات في الإحصاء ونظرية الاحتمالات لفهم كيفية تفاعل متغيرتين مع بعضهما البعض.[1]

يتميز المتغير العشوائي الثنائي او المزدوج بالخصائص التالية:[2]

-النتائج الزوجية: يتم تمثيل المتغيرات العشوائية الثنائية بواسطة ازواج من النتائج "(X,Y)"، حيث يمثل كل من "X,Y " متغيرة عشوائية مختلفة؛

-التوزيع المشترك: يتم وصف العلاقة بين المتغيرتين من خلال توزيع احتمالي مشترك، والذي يحدد احتمالات حدوث كل زوج من النتائج؛

-التطبيقات: تستخدم المتغيرات العشوائية الثنائية في العديد من المجالات، مثل تحليل البيانات، حيث يمكن ان تساعد في فهم العلاقات بين المتغيرات، مثل العلاقة بين الدخل والتعليم او بين العوامل البيئية والامراض؛

-يمكن ان تكون المتغيرات العشوائية الثنائية متصلة (مستمرة)، او عشوائية منفصلة (متقطعة).

ان من اهم استعمالات المتغيرات العشوائية المشتركة هي:

-تستعمل في إدارة المخاطر المتعلقة بالمحافظ الاستثمارية، بحيث انه في اغلب الاحياء المستثمرين في السوق المالي لا يهتمون بعائد أصل مالي واحد فقط، بال بالتوزيع المشترك لعوائد عدة أصول مالية؛

-وتستخدم أيضا بالتنبؤ بإنفاق المستهلك على مختلف حاجياته على السلع التي تشبعها في حالة زيادة دخله، هذا ما يعرف بدراسة سلوك المستهلك؛

-وتستخدمها البنوك في تقدير احتمالات التعثر المزدوج لتقييم المخاطر الائتمانية لها؛

-وتستخدم في مجال التامين لدراسة احتمالات وقوع الحوادث المترابطة، مثل تقدير العلاقة بين عدد الحوادث واعمار السائقين الذين قاموا بها، لتقدير الأقساط التأمينية العادلة.

---

[1]-مزواغي جيلالي، "**إحصاء 3**"، مرجع سبق ذكره، ص97.

[2]-نفس المرجع السابق، ص97.

2-المتغيرات العشوائية المزدوجة المنفصلة:[1]

بفرض ان "X, Y" متغيرين عشوائيين منفصلتين، حيث ان المتغير "X" يأخذ القيم المنفصلة:

$$X: x=x_1,x_2, \ldots, x_r$$

والمتغير "Y" يأخذ القيم المنفصلة:

$$Y: y=y_1,y_2, \ldots, y_c$$

فان دالة الاحتمال المشترك والتي تبين احتمال حدوث زوج القيم "(x, y)" معا تكتب على الصورة التالية:

$$\mathrm{P}\left(x=x_i,y=y_j\right)=f(x_i,y_j)$$

وتتصف هذه الدالة بالخاصيتين التاليتين:

$$- \ 0<f(x_i,y_j)<1 \ ;$$

$$- \ \sum_{j=1}^{c}\sum_{i=1}^{r} f(x_i,y_j)=1$$

ويمكن من دالة الاحتمال المشترك تكوين جدول التوزيع الاحتمالي المزدوج للمتغيرتين العشوائيتين المنفصلتين "X, Y"، وحساب احتمالات حدوث قيم المتغيرين معا، والتغاير بين المتغيرتين، كما يمكن استنتاج الخصائص الإحصائية التالية:

-دوال الاحتمال الهامشية واستخدامها في حساب متوسط وتباين التوزيع الهامشي؛

-التغاير بين المتغيرتين؛

-حساب معامل الارتباط البسيط بين المتغيرتين؛

-دالة الاحتمال الشرطية، واستخداماتها في حساب متوسط وتباين التوزيع الشرطي؛

-استنتاج دالة الانحدار من متوسط التوزيع الشرطي.

[1]-بلوطار مهدي، "**إحصاء 3**"، دروس عبر الخط، جامعة قالمة، ص1.

نستطيع ان نعطي نموذج مبسط لجدول الاقتران الاحتمالي كما يلي:[1]

	y_1	y_2	y_3	...	y_c	Sum
x_1	$f(x_1,y_1)$	$f(x_1,y_2)$	$f(x_1,y_3)$	...	$f(x_1,y_c)$	$W(x_1)$
x_2	$f(x_2,y_1)$	$f(x_2,y_2)$	$f(x_2,y_3)$	...	$f(x_2,y_c)$	$W(x_2)$
x_3	$f(x_3,y_1)$	$f(x_3,y_2)$	$f(x_3,y_3)$	...	$f(x_3,y_c)$	$W(x_3)$
...	...	...	...	...	...	...
x_r	$f(x_r,y_1)$	$f(x_r,y_2)$	$f(x_r,y_3)$	...	$f(x_r,y_c)$	$W(x_r)$
Sum	$G(y_1)$	$G(y_2)$	$G(y_{31})$	...	$G(y_c)$	1

3-المتغيرات العشوائية المزدوجة المتصلة:[2]

ليكن "X, Y" متغير عشوائي ذو بعدين الذي نفترض ان جميع قيمه في مجموعة جزئية غير محدودة من الفضاء الإقليدي، فان مثل هذا المتغير يسمى متغير عشوائي مستمر ذو بعدين، ومجموع القيم التي يأخذها تسمى مدى الفضاء ويرمز لها بالرمز:

$R_{X, Y}$

ان دالة التوزيع المزدوجة (ثنائية المتغيرات) تعطي احتمالية ان يأخذ "(x, y)" قيما في المنطقة المحددة "A" وهي تعطى بالعلاقة التالية:

$$\mathrm{P}(a<X<b, c<Y<d)=\int_a^b \int_c^d f_{x,y}(x,y)dxdy$$

ان دالة الاحتمال المشترك تكون دائما غير سلبية، والتكامل المزدوج يعطي قيمة واحد (1). أي:

- $f_{x,y}(x,y)\geq 0$;
- $\int_a^b \int_c^d f_{x,y}(x,y)dxdy = 1$;
- $f_y(y) = \int_{-\infty}^{+\infty} f_{x,y}(x,y)dx.$

[1]-مزواغي جيلالي، "**إحصاء 3**"، مرجع سبق ذكره، ص97.

[2]-نفس المرجع السابق، ص99.

4-المميزات العددية للمتغيرات العشوائية المزدوجة:

ان المميزات العددية للمتغيرات العشوائية المزدوجة تتمثل في:

-التوقع الرياضي:

التوقع الرياضي للمتغيرات العشوائية المزدوجة المنفصلة هو:

$$E(X)=\mu_x=\sum_x\sum_y xP(x,y)$$

$$E(Y)=\mu_y=\sum_x\sum_y yP(x,y)$$

التوقع الرياضي للمتغيرات العشوائية المزدوجة المتصلة هو:

$$E(X)=\mu_x=\int\int xf(x,y)dxdy$$

$$E(Y)=\mu_y=\int\int yf(x,y)dxdy$$

-التباين:

التباين للمتغيرات العشوائية المزدوجة المنفصلة هو:

$$\sigma_X^2=\sum_x\sum_y (x-E(x)^2)P(x,y)$$

$$\sigma_Y^2=\sum_x\sum_y (y-E(y)^2)P(x,y)$$

التباين للمتغيرات العشوائية المزدوجة المتصلة هو:

$$\sigma_X^2=\iint (x-E(x)^2)f(x,y)$$

$$\sigma_Y^2=\iint (y-E(y)^2)f(x,y)$$

-التباين المشترك:

التباين المشترك للمتغيرات العشوائية المزدوجة المنفصلة هو:

$$\mathrm{Cov}(x,y)=\sigma_{xy}=\sum_i\sum_j (x_i-\mu_X)(y_j-\mu_y)P(x,y)$$

$$\mathrm{Cov}(x,y)=\sigma_{xy}=E(x,y)-E(x)E(y)$$

التباين المشترك للمتغيرات العشوائية المزدوجة المتصلة هو:

$$\mathrm{Cov}(x,y)=\sigma_{xy}=\int R\int (x_i-\mu_X)(y_j-\mu_y)f(x,y)$$

-معامل الارتباط:

ان معامل الارتباط يقيس قوة العلاقة بين المتغيرات وصيغته تكتب كما يلي:

$$\rho(x,y)=\frac{cov(x,y)}{\sigma_x\sigma_y}$$

حيث ان:

$$0\leq\rho\leq 1$$

علما ان إذا كان:

$$\rho(x,y)=0$$

هذا يعني ان المتغيرتين العشوائيتين "X, Y" مستقلتين.

6-تمارين محلولة:[1]

التمرين الأول

ليكن لدينا المتغيرتين العشوائيتين "X, Y" معرفتين كما يلي:

-المتغيرة العشوائية "X" معرفة كما يلي:

-إذا كان الشخص يدخن أي "$X=1$"؛

-إذا كان الشخص لا يدخن "$X=0$".

-المتغيرة العشوائية "Y" معرفة كما يلي:

-الشخص مصاب بمرض "$Y=1$"؛

-الشخص غير مصاب بمرض "$Y=0$".

حيث يكون جدول قانون التوزيع الاحتمالي المشترك كما يلي:

	$Y=0$	$Y=1$
$X=0$	0.40	0.02
$X=1$	0.03	0.55

مما سبق اوجد احتمال ان يكون الشخص لا يدخن وغير مصاب بمرض والشخص لا يدخن ومصاب بمرض.

$-P(X=0,Y=0)=h(0,0)$

$$P(X=0,Y=0)=0.40$$

$-P(X=0,Y=1)=h(0,1)$

$$P(X=0,Y=1)=0.02$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

التمرين الثاني

بصندوق ثلاث كرات حمراء وكرتين بيضاء، نقوم بالسحب عشوائيا كرتين بدون ارجاع، فليكن:

-"X": هي متغيرة عشوائية تمثل عدد الكرات الحمراء المسحوبة؛

-"Y": هي متغيرة عشوائية تمثل عدد الكرات البيضاء المسحوبة.

1-انشاء جدول التوزيع الاحتمالي المشترك لـ "X, Y"؛

2-احسب التوقع الرياضي المشترك لـ "X, Y".

1-انشاء جدول التوزيع الاحتمالي المشترك لـ "X, Y":

بما ان السحب يتم بكرتين فقط، فان الحالات الممكنة للزوج "(X,Y)" هي كالاتي:

-كرتين حمراء و0 كرة بيضاء: (2، 0)؛

-كرة حمراء وكرة بيضاء: (1، 1)؛

-0 كرة حمراء وكرتين بيضاء: (0، 2).

ان عدد طرق سحب كرتين من أصل 5 هي:

$$\binom{5}{2}=10$$

$$P(X=2,Y=0)=\frac{\binom{3}{2}\times\binom{2}{0}}{10}$$

$$P(X=2,Y=0)=\frac{3\times1}{10}$$

$$\boldsymbol{P(X=2,Y=0)=0.3}$$

$$P(X=1,Y=1)=\frac{\binom{3}{1}\times\binom{2}{1}}{10}$$

$$P(X=1,Y=1)=\frac{3\times2}{10}$$

$$\boldsymbol{P(X=1,Y=1)=0.6}$$

$$P(X=0,Y=2)=\frac{\binom{3}{0}\times\binom{2}{2}}{10}$$

$$P(X=0,Y=2)=\frac{1\times1}{10}$$

$$\mathbf{P(X=0,Y=2)=0.1}$$

ومنه جدول التوزيع الاحتمالي المشترك لـ "X, Y" هو كما يلي:

	$Y=0$	$Y=1$	$Y=2$	Sum
$X=0$	0	0	0.1	0.1
$X=1$	0	0.6	0	0.6
$X=2$	0.3	0	0	0.3
Sum	0.3	0.6	0.1	1

2-احسب التوقع الرياضي المشترك لـ "X, Y":

$$E(X+Y)=\sum\sum(X+Y)P(X,Y)$$

$$E(X+Y)=(2+0)(0.3)+(1+1)(0.6)+(0+2)(0.1)$$

$$E(X+Y)=(2)(0.3)+(2)(0.6)+(2)(0.1)$$

$$E(X+Y)=0.6+1.2+0.2$$

$$\mathbf{E(X+Y)=2}$$

التمرين الثالث

ليكن لدينا التوزيع المشترك الاحتمالي في الجدول التالي:

	Y=-3	***Y=2***	***Y=4***	***Sum***
X=1	*0.1*	*0.2*	*0.2*	*0.5*
X=3	*0.3*	*0.1*	*0.1*	*0.5*
Sum	*0.4*	*0.3*	*0.3*	***1***

-احسب التباين المشترك؛

-احسب معامل الارتباط، هل المتغيرين "X, Y" مستقلان؟

$$\text{cov}(x,y)=\sigma_{xy}=\sum_{i}\sum_{j} x_i y_j P(x,y)-\mu_x\mu_y$$

$$\mu_x=\sum x_i P(x_i)=1(0.5)+3(0.5)=2$$

$$\mu_y=\sum y_j P\left(y_{ji}\right)=-3(0.4)+2(0.3)+4(0.3)=0.6$$

ومنه نجد:

$$cov(x,y)=\sigma_{xy}=1(-3)(0.1)+1(2)(0.2)+1(4)(0.2)+3(-3)(0.3)+3(2)(0.1)+3(4)(0.1)-$$

$$\boldsymbol{cov(x,y)=-1.2}$$

$$\sigma_x^2=\sum x_i^2 P(x_i)-\mu_x^2$$

$$\sigma_x^2=\left(1(0.5)+9(0.5)\right)-(2)^2$$

$$\boldsymbol{\sigma_x^2=1}$$

$$\sigma_y^2=\sum y_i^2 P(y_i)-\mu_y^2$$

$$\sigma_y^2=(9(0.4)+4(0.3)+16(0.3))-(0.6)^2$$

$$\boldsymbol{\sigma_y^2=9.24}$$

$$\rho(x,y)=\frac{cov(x,y)}{\sigma_x\sigma_y}$$

$$\rho(x,y)=\frac{-1.2}{\sqrt{1}\sqrt{9.24}}$$

$$\rho(x,y)=-0.4$$

ان المتغيرتين العشوائيتين "X, Y" غير مستقلتين لان معامل الارتباط بينهما لا يساوي الصفر (0).

التمرين الرابع

ليكن لدينا جدول التوزيع الاحتمالي المشترك للزوج (X,Y):

	Y=5	Y=10	Y=15	Sum
X=2	0	0	0.06	0.06
X=4	0.02	0.04	0.1	0.16
X=6	0.08	0.1	0.12	0.3
X=8	0.1	0.14	0.02	0.24
X=10	0.16	0.06	0	0.22
Sum	0.36	0.34	0.30	1

احسب الاحتمالات المشتركة التالية:

-$P(x<8,y<15)$

-$P(4\leq x<10,5<y<15)$

-$P(x<8,y<15)=\sum_{i=1}^{3}\sum_{j=1}^{2}P_{ij}$

$$\sum_{i=1}^{3}(P_{i1}+P_{i2})=(P_{11}+P_{21}+P_{31})+(P_{12}+P_{22}+P_{32})$$

$$\sum_{i=1}^{3}(P_{i1}+P_{i2})=(0+0+0.06)+(0.02+0.04+0.1)$$

$$\sum_{i=1}^{3}(P_{i1}+P_{i2})=0.52$$

-P(4≤x<10,5<y<15)=P(x_2≤x<x_5,y_1<y<y_3)

$$P(4\leq x<10,5<y<15)=0.02+0.08+0.1+0.04+0.1+0.14$$

$$P(4\leq x<10,5<y<15)=0.48$$

المحاضرة الرابعة عشر

دوال المتغيرات العشوائية المزدوجة

المحاضرة الرابعة عشر: دوال المتغيرات العشوائية المزدوجة

ان:

– دوال المتغيرات العشوائية المزدوجة "*Functions Bivariate random variables*": تطور مفهومها مع تطور نظرية الاحتمالات واستعمالاتها، فلقد بدا المصطلح بالتطور والتأسيس بظهور مفهوم المتغيرات الفردية المنفصلة منها والمتصلة، ليتوسع المفهوم ويشمل دوال المتغيرات المزدوجة، فهذه الدوال هي عبارة عن تطور وتفسير وصياغة رياضية واحصائية للظواهر المعقدة والمتشابكة، ويعتبر كل من "يوهان كارل فريديريك غوس" " *Johann Carl Friedrich Gauss*" (1777-1855) و جاكوب برنولي" "*Jakob Bernoulli* " (1654-1705) من رواد الاعمال الإحصائية والرياضية، التي أسهمت في تأسيس دوال المتغيرات العشوائية المزدوجة، من خلال وضع الأسس للاحتمال الشرطي والتوزيعات المشتركة بين المتغيرات العشوائية، لفهم العلاقة بينها، وهي تعتبر مهمة في الاقتصاد والفيزياء لتفسير الظواهر المتعددة المتغيرات.

1-دوال المتغيرات العشوائية المزدوجة:

يعرف التوزيع المشترك بانه دالة احتمالية تجمع بين عدة متغيرات عشوائية في آن واحد، فعلى فرض ان "X_1,X_2" متغيران عشوائيان عندئذ فان النموذج الرياضي الاحتمالي الذي يعبر عن سلوك هذين المتغيرين معا يسمى "التوزيع الاحتمالي المشترك" للمتغيرين "X_1,X_2"، او "التوزيع الاحتمالي لمتغيرين" "*Bivariate distribution*". كما هو الحال لدى دراسة العلاقة بين الكمية المعروضة من سلعة معينة "X_1"، وسعر الوحدة الواحدة منها "X_2"، وبشكل أكثر عمومية إذا كانت "$X_1,X_2,\dots,X_k$" متغيرات عشوائية عندئذ فان النموذج الرياضي الاحتمال الذي يعبر عن سلوك هذه المتغيرات مجتمعة يسمى "التوزيع الاحتمالي المشترك" للمتغيرات "$X_1,X_2,\dots,X_k$"، او "التوزيع الاحتمالي لعدة متغيرات" "*Multivariate distribution*". ومن وجهة نظر احتمالية يمكن تحديد نوعين رئيسيين من التوزيعات المشتركة، استنادا الى نوع المتغيرات المتضمنة في التوزيع المشترك من حيث كونها متغيرات منفصلة ام متصلة.[1] وهي على التوالي:

-دوال المتغيرات العشوائية المزدوجة المنفصلة؛

-دوال المتغيرات العشوائية المزدوجة المتصلة.

2-دوال المتغيرات العشوائية المزدوجة المنفصلة:

لنفرض ان:[2]

X_1,X_2

متغيران عشوائيان من النوع المنفصل، عندئذ فان دالة الكثافة الاحتمالية التي تعبر عن سلوك هذين المتغيرين مع هي:

$$P(X_1,X_2)=P(X_1=x_1,X_2=x_2)$$

وهذا يعني:

$$P(X_1,X_2)=P(X_1=x_1 \cap X_2=x_2)$$

أي احتمال وقوع الحادثة:

$$E=\{X_1=x_1 \cap X_2=x_2\}$$

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص149.

[2]-نفس المرجع السابق، ص150.

وبشكل عام إذا كانت "$X_1, X_2, \dots, X_k$" متغيرات عشوائية من النوع المنفصل، فإذن دالة الكثافة الاحتمالية التي تعبر عن سلوك هذه المتغيرات مجتمعة هي:

$$P(X_1, X_2, \dots, X_k) = P(X_1 = x_1 \cap X_2 = x_2 \dots \cap X_k = x_k)$$

أي احتمال وقوع الحادثة:

$$E = \{X_1 = x_1 \cap X_2 = x_2 \dots \cap X_k = x_k\}$$

ان دالة التوزيع المشترك يجب ان تحقق الشروط التالية التي تسمح لنا اعتبارها دالة كثافة احتمالية مشتركة وهي:

-ان دالة التوزيع المشترك دالة وحيدة القيمة عند اية قيمة مخصصة لمتغيرات التوزيع مثل "$X_1, X_2, \dots, X_k$"؛

-ان دالة التوزيع المشترك دالة غير سالبة، كونها تعبر عن احتمال وقوع حادثة مثل "E"، أي ان:

$$P(X_1, X_2, \dots, X_k) \geq 0$$

-ان مجموع الكتل الاحتمالية المشتركة المقترنة بالقيم الممكنة الى "$X_1, X_2, \dots, X_k$" يجب ان يكون مساويا للواحد، أي ان:

$$\sum P(X_1, X_2, \dots, X_k) = 1$$

3-دوال المتغيرات العشوائية المزدوجة المتصلة:

نفترض ان "X_1, X_2" متغيران عشوائيان من النوع المتصل، عندئذ فان دالة الكثافة الاحتمالية المشتركة التي تعبر عن سلوك هذين المتغيرين معا هي "$f(x_1, x_2)$". ان "$f(X_1 = x_1, X_2 = x_2)$" لا تعبر هنا عن قيمة احتمالية بل تعبر عن قيمة الدالة "f" عند قيمة مخصصة الى "X_1, X_2" مثل "(x_1, x_2)".[1]

وليكن "X, Y" متغير عشوائي ذو بعدين الذي نفترض ان جميع قيمه في مجموعة جزئية غير محدودة من الفضاء الاقليدي، تعطى دالة التوزيع التراكمي لمتغير عشوائي ثنائي متصل بالعلاقة التالية:[2]

$$F_{x,y}(x,y) = P(X < x, Y < y)$$

$$F_{x,y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(x,y)}(t_1, t_2) dt_1 dt_2$$

وما تجدر الإشارة اليه ان:

---

[1]-نفس المرجع السابق، ص102.

[2]-مزواغي جيلالي، "**إحصاء 3**"، مرجع سبق ذكره، ص102.

$$F_{x,y}(x,y)=\frac{\partial^2}{\partial x\partial y}F_{x,y}(x,y)$$

بشكل عام إذا كانت "$X_1,X_2,\dots,X_k$" متغيرات عشوائية من النوع المتصل، فان الدالة "$f(x_1,x_2,\dots,x_k)$"، تسمى دالة الكثافة الاحتمالية المشتركة التي تعبر عن سلوك المتغيرات "$X_1,X_2,\dots,X_k$" مجتمعة، وهذا النوع من الدوال يجب ان يحقق الخصائص التالية التي تسمح لنا اعتبار "f" دالة كثافة احتمالية مشتركة وهي:[1]

-ان الدالة "f" دالة وحيدة القيمة عند اية قيمة مخصصة الى $(X_1,X_2,\dots,X_k)$ مثل $(x_1,x_2,\dots,x_k)$؛

-ان الدالة "f" دالة غير سالبة عند اية قيمة مخصصة الى $(X_1,X_2,\dots,X_k)$ مثل $(x_1,x_2,\dots,x_k)$؛

-انه وبشكل عام:

$$\int\int\dots\int f(x_1,x_2,\dots,x_k)dx_k dx_{k-1}\dots dx_1$$

4-التمثيل البياني لدوال المتغيرات العشوائية المزدوجة:

ان التمثيل البياني لدوال المتغيرات العشوائية المزدوجة يهدف لإظهار وتوضيح العلاقة التي تربط متغيرين او أكثر، وشكل التمثيل البياني لدوال المتغيرات العشوائية المزدوجة يعتمد على نوع المتغيرات العشوائية، ان كانت منفصلة او متصلة كما يلي:

4-1-التمثيل البياني لدوال المتغيرات المزدوجة المنفصلة:

التمثيل البياني لدوال المتغيرات المزدوجة المنفصلة يهدف لتوضيح العلاقة بين المتغيرين العشوائيين "X,Y"، واحتمالات وقوعها او حدوثها في ان واحد.

ان التمثيل البياني لدوال المتغيرات العشوائية المنفصلة تعتمد على ثلاثة محاور كما يلي:

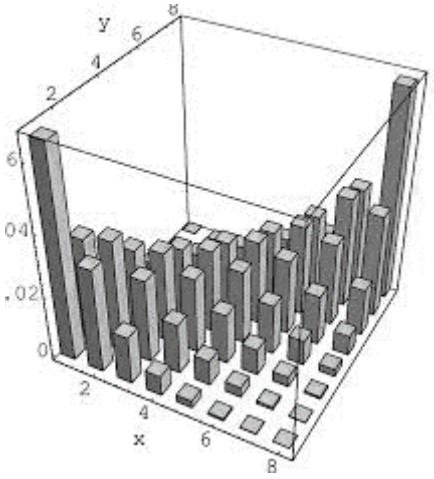
-المحور الافقي الاول يخصص للمتغير العشوائي "X"، والذي يمثل قيم المتغير العشوائي الأول؛

-المحور الافقي الثاني يخصص للمتغير العشوائي "Y"، والذي يمثل قيم المتغير العشوائي الثاني؛

-المحور العمودي يخصص للاحتمال "Z"، والذي يمثل قيم دالة الاحتمال المشتركة للمتغيرين العشوائيين "X,Y".

[1]-أمير حنا هرمز، "**الإحصاء الرياضي**"، مرجع سبق ذكره، ص102.

نستطيع ان نمثل الاحتمال على شكل أعمدة لكل زوج من المتغيرين العشوائيين "X,Y". ونستطيع ان نمثل دوال المتغيرات العشوائية المنفصلة كما يلي:

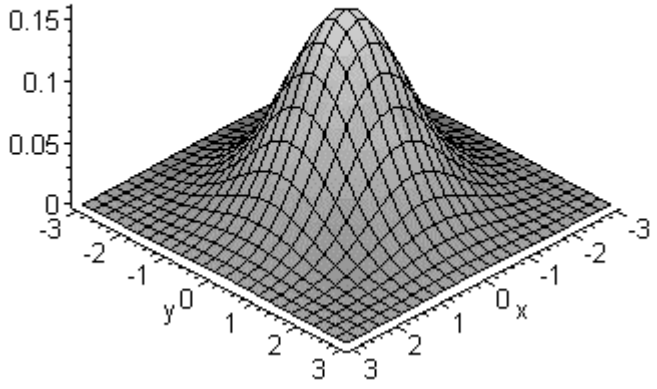


4-1-التمثيل البياني لدوال المتغيرات المزدوجة المتصلة:

التمثيل البياني لدوال المتغيرات العشوائية المزدوجة المتصلة يهدف هو أيضا لتمثيل العلاقة بين متغيرين عشوائيين "X,Y"، ولكن بصورة مستمرة. ويتم تمثيل البياني لهذه الدوال عن طريق دالة الكثافة الاحتمالية المزدوجة (المشتركة) بين المتغيرات العشوائية.

التمثيل البياني لدوال المتغيرات العشوائية المزدوجة المتصلة تتشابه من حيث المحاور، فهي تتشابه من حيث عدد المحاور والمتغيرات التي تمثلها، فالمحور الافقي الأول يخص المتغير العشوائي "X"، اما المحور الافقي الثاني فيخصص للمتغير العشوائي "Y"، اما المحور الثالث العمودي يخصص للاحتمالات، ووجه الاختلاف بينهما هو في قيم الاحتمالات "Z" فهي مستمرة في هذه الحالة.

يستطيع ان يكون التمثيل البياني لدوال المتغيرات العشوائية المزدوجة المتصلة كما يلي:



5-تمارين محلولة:[1]

التمرين الأول

لتكن دالة التوزيع الاحتمالية المزدوجة للمتغيرين العشوائيين "X,Y" هي:

$$f(x,y)=\begin{cases}\left(\frac{3}{y}\right)\left(\frac{1}{4}\right)^{x+1} & , x=1,2,\ldots,\infty;y=1,2,3\\ 0 & or\end{cases}$$

المطلوب:

ايجاد دالة التوزيع التجميعية المزدوجة.

لإيجاد دالة التوزيع التجميعية المزدوجة نستخدم العلاقة التالية:

$$F_{x,y}(x,y)=P(X<x,Y<y)$$

$$F_{x,y}(x,y)=\sum_{t_1\in IR(X)t_1\leq x}\sum_{t_2\in IR(Y)t_2\leq y} f_{(x,y)}(t_1,t_2)$$

حيث ان دالة التوزيع الاحتمالية المزدوجة المعطاة هي:

$$f(x,y)=\begin{cases}\left(\frac{3}{y}\right)\left(\frac{1}{4}\right)^{x+1} & , x=1,2,\ldots,\infty;y=1,2,3\\ 0 & or\end{cases}$$

تحديد حدود الجمع:

-"X" يأخذ القيم من 1 الى "t_1"؛

-"Y" يأخذ القيم من 1 الى "t_2".

بما ان:

$$\begin{cases}x>1\\ y>1\end{cases}$$

[1]-في اعداد التمرينات تم الاعتماد على مجموعة من المراجع.

فان:

$$F_{x,y}(x,y)=\sum_{t_1=1}\sum_{t_2=1} f_{(x,y)}(t_1,t_2)$$

$$F_{x,y}(x,y)=\sum_{t_1=1}\sum_{t_2=1} 3t_2\left(\frac{1}{4}\right)^{t_1+1}$$

التمرين الثاني

بمعطيات التمرين السابق، قم بحساب التالي:

$F_{x,y}(2,2)$

لدينا:

$$F_{x,y}(x,y)=\sum_{t_1=1}\sum_{t_2=1} 3t_2\left(\frac{1}{4}\right)^{t_1+1}$$

فان:

$$F_{x,y}(2,2)=\sum_{t_1=1}^{2}\sum_{t_2=1}^{2} 3t_2\left(\frac{1}{4}\right)^{t_1+1}$$

-عندما:

$$\begin{cases} t_1=1 \\ t_2=1 \end{cases}$$

$$F_{x,y}(1,1)=3(1)\left(\frac{1}{4}\right)^{1+1}$$

$$F_{x,y}(1,1)=\frac{3}{16}$$

-عندما:

$$\begin{cases} t_1 = 1 \\ t_2 = 2 \end{cases}$$

$$F_{x,y}(1,1) = 3(2)(\frac{1}{4})^{1+1}$$

$$F_{x,y}(1,1) = \frac{6}{16}$$

-عندما:

$$\begin{cases} t_1 = 2 \\ t_2 = 1 \end{cases}$$

$$F_{x,y}(1,1) = 3(1)(\frac{1}{4})^{2+1}$$

$$F_{x,y}(1,1) = \frac{3}{64}$$

-عندما:

$$\begin{cases} t_1 = 2 \\ t_2 = 2 \end{cases}$$

$$F_{x,y}(1,1) = 3(2)(\frac{1}{4})^{2+1}$$

$$F_{x,y}(1,1) = \frac{6}{64}$$

$$F_{x,y}(2,2) = \sum_{t_1=1}^{2}\sum_{t_2=1}^{2} 3t_2(\frac{1}{4})^{t_1+1}$$

$$F_{x,y}(2,2) = \frac{3}{16} + \frac{6}{16} + \frac{3}{64} + \frac{6}{64}$$

$$\mathbf{F_{x,y}(2,2) = \frac{45}{64}}$$

قائمة المراجع

قائمة المراجع

-أمير حنا هرمز، "**الإحصاء الرياضي**"، دار ابن الاثير للطباعة والنشر في جامعة الموصل، 1990.

-باسل يونس ذنون، "**الاحتمالية والمتغيرات العشوائية**"، دار الكتب للطباعة والنشر، الموصل، 1991.

-بلوطار مهدي، "**إحصاء 3**"، دروس عبر الخط، جامعة قالمة.

-بوزنورة اسماء، "**الإحصاء (02) –دروس وتمارين–**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 03، 2024/2023.

-بولعباس مختار، "**محاضرات في الإحصاء 2 مدعومة بتمارين محلولة**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة تيارت، 2023/2022.

-جنيدي مراد، "**محاضرات في الإحصاء 3**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 3، 2021/2020.

-حمبلي زهير، "**دروس عبر الخط في مقياس الإحصاء 03**"، جامعة ام البواقي.

-رامز قدسية، "**الاحتمالات والاحصاء**"، منشورات الجامعة الافتراضية السورية، 2018.

-طويطو محمد، "**مطبوعة خاصة بمقياس الإحصاء 03**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 03، 2017/2016.

-عبدلي ادريس، "**الإحصاء 03**"، دروس عبر الخط في مقياس، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة البليدة 02، 2024/2023.

-علاء الدين القبانجي وحسام حمامة كمرجي، "**الاحتمال والاحصاء**"، منشورات جامعة دمشق، 2012.

-علي عبد السلام العماري وعلي حسين العجيلي، "**الإحصاء والاحتمالات النظرية والتطبيق**"، دار الحكمة، 2000.

-مزواغي جيلالي، "**إحصاء 3**"، مطبوعة بيداغوجية، جامعة غيليزان، 2025/2024.

-همال فريدة، "**الإحصاء 3 دروس وتمارين**"، مطبوعة بيداغوجية، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، جامعة الجزائر 3، 2022/2021.