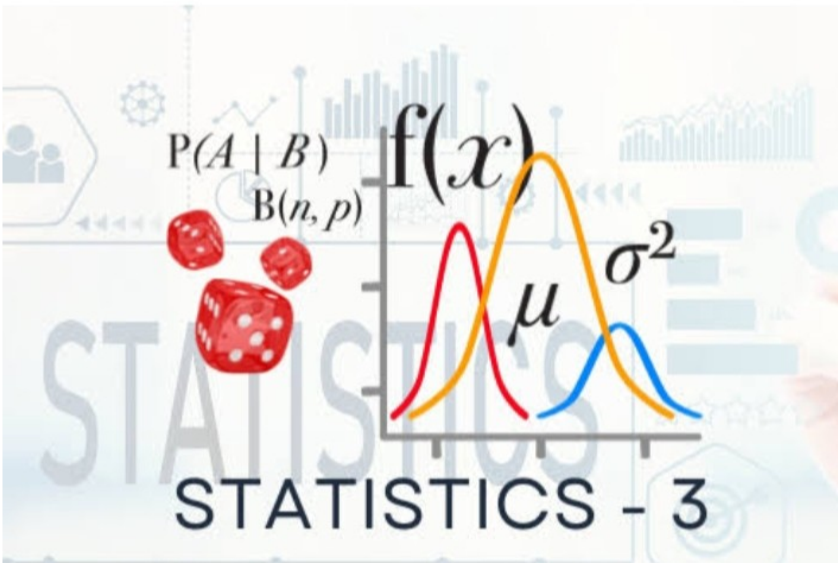
كلية العلوم الاقتصادية والتجارية وعلوم التسيير

أهم قوانين التوزيعات الاحتمالية للمتغيرات المستمرة

الإحصاء 3 -

hamida.hassini@univ-



جامعة تلمسان

08/03/2024

مفتاح المصطلحات

مدخل القاموس	
مرجع عام	

قائمة المحتويات

وحدة

• نهدف من خلال دراسة هذا المحور إلى:

- التعرف على أهم التوزيعات الاحتمالية المستمرة مثل التوزيع الطبيعي، التوزيع الأسي، توزيع كاي تربيع ، وتوزيع ستيودنت.
- دراسة المفاهيم الأساسية مثل المتغيرات العشوائية المستمرة، والدالة الاحتمالية الكثافة (PDF)، والدالة التوزيع التراكمي (CDF).
- فهم خصائص هذه التوزيعات
- استخدام التوزيعات المستمرة في نمذجة الظواهر الإقتصادية ، وتحليل البيانات.

مقدمة

إن دراسة التوزيعات الاحتمالية للمتغيرات المستمرة هي أحد الجوانب الرئيسية في الإحصاء الرياضي ، حيث توفر إطارًا نظريًا لفهم كيفية توزيع القيم المحتملة للمتغيرات العشوائية المستمرة. و نهدف من خلال دراسة هذا المحور إلى استكشاف النماذج الرياضية التي تصف التوزيع الاحتمالي لهذه المتغيرات، مما يتيح لنا التنبؤ بالسلوك الاحتمالي للأنظمة والظواهر المختلفة.

ويتضمن هذا المحور التعرف على أهم التوزيعات الاحتمالية المستمرة مثل التوزيع الطبيعي، والتوزيع الأسي، وتوزيع كاي تربيع، والتوزيع فيشر، والتوزيع ستيودنت. من خلال فهم هذه التوزيعات وخصائصها

المحور الثاني: أهم قوانين التوزيعات الاحتمالية للمتغيرات المستمرة

أ. التوزيع المنتظم (Uniform Distribution)

هو أحد التوزيعات الاحتمالية البسيطة، حيث لدى كل قيمة من القيم داخل مجال معين نفس الإحتمال .
بمعنى آخر، إذا كان المتغير العشوائي المستمر X يتبع توزيعًا منتظمًا على الفترة[a-b] فإن احتمال وقوع X في أي جزء من هذه الفترة هو نفسه، بشرط أن تكون الأطوال متساوية.

التعريف الرياضي للتوزيع المنتظم على الفترة [a-b]هو كالتالي:

دالة الكثافة الاحتمالية (**PDF**):

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{إذا كان } a \leq x \leq b, \\ 0 & \text{خلاف ذلك.} \end{cases}$$

دالة التوزيع التراكمي (**CDF**):

$$F(x) = \begin{cases} 0 & \text{إذا كان } x < a, \\ \frac{x-a}{b-a} & \text{إذا كان } a \leq x \leq b, \\ 1 & \text{إذا كان } x > b. \end{cases}$$

الخصائص الرئيسية للتوزيع المنتظم هي:

المتوسط:

$$\mu = \frac{a+b}{2}$$

التباين:

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

ب. التوزيع الأسي (Exponential Distribution)

عادة ما يستخدم التوزيع الأسي في مسائل متعلقة بقياس الزمن. مثل مدة خدمة شباك البريد، مدة مكالمة هاتفية، مدة تفريغ باخرة شحن، مدة تصليح آلة، مدة انتظار زبون قبل الحصول على الخدمة...

ومن الضروري فهم الآتي: كقاعدة عامة يستخدم التوزيع الأسي لتمثيل مدة حياة ظاهرة ما إذا كان لها متوسط ثابت(λ/1) وكانت هذه الظاهرة لا تخضع للتقادم (vieillissement) أي أن مدة حياة الظاهرة بعد لحظة ما T لا تتبع اللحظة T-1؛ أي لا تتأثر بالمدة التي دامتها الظاهرة من قبل. مثلا قد نستبعد استخدام التوزيع الأسي لتمثيل مدة حياة آلة عاملة قبل تعطلها لأن احتمال تعطلها في لحظة ما ليس مستقلا عن المدة التي عملتها الآلة من قبل، كذلك الأمر بالنسبة لمدة حياة الإنسان.

نشير أخيرا إلى أن للتوزيع الأسي علاقة بتوزيع بواسون، فإذا كان وقوع أحداث ما يتبع هذا التوزيع، فإن المدة بين وقوع حدثين تتبع التوزيع الأسي.

كمثال على ذلك، إذا كان وصول الزبائن إلى مركز خدمة ما يتبع توزيع بواسون فإن المدة الزمنية بين وصول زبون "أ" والزبون الموالي (ب) تتبع التوزيع الأسي. و تتبين هذه العلاقة عند استنتاج صيغة القانون الأسي.

دالة الكثافة الاحتمالية (**PDF**):

$$f(x;\lambda)=\begin{cases}\lambda e^{-\lambda x} & \text{if } x\geq 0,\\ 0 & \text{if } x<0,\end{cases}$$

λ هو معلمة المعدل (rate parameter) للتوزيع، ويمثل متوسط عدد الأحداث في وحدة الزمن.

دالة التوزيع التراكمي (**CDF**):

$$F(x;\lambda)=\begin{cases}1-e^{-\lambda x} & \text{if } x\geq 0,\\ 0 & \text{if } x<0.\end{cases}$$

تطبيقات التوزيع الأسي تشمل:1[3][3][1]

نظم الطوابير: نمذجة الوقت بين وصول العملاء إلى طابور الخدمة.

الاعتمادية والصيانة: تقدير وقت الفشل للأجهزة والمكونات.

الاقتصاد: التحليل الزمني بين الصفقات المالية أو الأحداث الاقتصادية.

- فهم التوزيع الأسي وخصائصه يساعد في تحليل وتفسير البيانات الزمنية التي تتبع نمط وصول الأحداث عشوائيا.

الخصائص الرئيسية للتوزيع الأسي :

المتوسط (التوقع الرياضي) والتباين:8[8]

$$\mu=1/\lambda \quad ، \quad \sigma^2=1/\lambda^2$$

پ. تمرين

[17 ص 1 حل رقم]

تمرين

إذا كان لديك المتغير x يتبع التوزيع المنتظم للفترة مابين 0 و 10، ما هو احتمال أن تكون القيمة المتغير بين 3 و 7؟

تمرين

التوزيع الأسي هو توزيع احتمالي مستمر يستخدم لنمذجة الفترات الزمنية بين الأحداث التي تحدث

بشكل ثابت ومستقل.

تمرين

إذا كان متوسط الوقت بين وصول العملاء إلى متجر ما هو 10 دقائق، ما هو قيمة λ للتوزيع الأسي في هذه الحالة؟

- ☐ 0.1 عميل لكل دقيقة
- ☐ 10 عميل لكل دقيقة

أجب بصحيح أو خطأ

يتبع المتغير العشوائي المستمر X الذي يمثل ظاهرة متعلقة بالزمن، التوزيع الأسي؟

ت. التوزيع الطبيعي (Normal Distribution) أو (Gaussian Distribution)

يعد التوزيع الطبيعي من أهم التوزيعات الاحتمالية شائعة الاستخدام لما له من خصائص تنطبق على نسبة كبيرة من الظواهر الطبيعية والاجتماعية والاقتصادية، فلو اخترنا بالصدفة مئة أو ألفا من المارين في شار ع ما وقسنا أطوالهم لوجدنا نسبة كبيرة منها قريبة من متوسط ما، ونسبة قليلة من طوال القامة ونسبة مقاربة لها من قصار القامة. ومثل هذا بالنسبة للأوزان. ولو مثلنا هذه البيانات في معلم متعامد متجانس لكان المنحنى الذي يمثل النسبة، أو ما يمكن أن نسميه الاحتمال، ذا شكل جرسي متماثل حول المتوسط وهي صفات التوزيع الطبيعي حيث μ و σ هما على التوالي التوقع والانحراف المعياري. ونكتب أن المتغير العشوائي x يتبع التوزيع الطبيعي على النحو التالي: $X \sim N(\mu,\sigma)$

دالة التوزيع (الدالة التجميعية) للتوزيع الطبيعي تكتب كما يلي: [4][3][3]4

$$F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-1/2\left(\frac{v-\mu}{\sigma}\right)^2} dv$$

المتغيرة المركزية أو المعيارية : تستخدم المتغيرة المعيارية $Z = (X-\mu)/\sigma$ لتكوين الجداول الإحصائية للاحتمالات:

$F(z) = P(Z \le z)$ أو $P(0 \le Z \le z)$،

حيث تسمح بكتابة الدالة F بدلالة مجهول واحد وهو Z بدلا من 3 مجاهيل x و μ و σ وذلك كما يلي:

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty}$$

و لقد تم باستخدام المتغيرة المعيارية Z لحساب الاحتمالات (المساحات) تحت المنحنى ومنها :

$P(-\sigma \le X \le \sigma) = P(-\sigma \le Z \le \sigma) = 0.6837$,

$P(-2\sigma \le X \le 2\sigma) = P(-2 \le Z \le 2) = 0.9544$,

$P(-3\sigma \le X \le 3\sigma) = P(-3 \le Z \le 3) = 0.9973$.

هذه القيم وغيرها متوفرة في الجداول الإحصائية التي نجدها في الكثير من المراجع، كما يمكن حسابها

باستخدام الحاسوب.
[3]3

ث. توزيع كاي تربيع $\chi 2$ distribution

توزيع ك 2 هو من أكثر التوزيعات استخداما في مجال اختبار الفروض بأنواعها، ويمكن تعريفه كما يلي:
لتكن $X1, X2, . . . Xv$ ، متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري $(\mu = 0, \sigma=1)$.
و المتغيرة $X = X1^2 + X2^2 + . . . + Xv^2$
لها دالة الكثافة التالية:

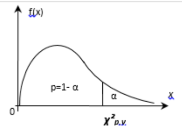
$$f(x)=\begin{cases}\dfrac{x^{(v/2)-1}e^{-x/2}}{2^{v/2}\Gamma(v/2)} & x>0\\ 0 & x\le 0\end{cases}$$

حيث أن : حيث $\Gamma(\alpha)$ هي الدالة قاما: $\alpha > 0$ $\quad \Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt$

وعليه نقول أن X تتبع التوزيع ك 2 ب v درجة حرية ونكتب $X \sim \chi^2 v$.
الدالة التجميعية $F\ (\chi^2)$ تكتب كما يلي:

$$P(X^2\le x)=\begin{cases}\dfrac{1}{2^{v/2}\Gamma(v/2)}\displaystyle\int_0^x u^{(v/2)-1}e^{-u/2}du & x\ge 0\\ 0 & x<0\end{cases}$$

خصائص توزيع ك 2
$E(X) = v,\ V(X) = 2v,\ M(t) = (1-2t)^{-v/2}$
دالة التوزيع ك 2 هي هي حالة خاصة من توزيع قاما بوضع $\alpha = v/2, \beta = 2$.
ويأخذ منحنى $f(x)$ شكله حسب قيمة الثابت v ونلاحظ من الرسم أن المنحنى يبتعد شيئا فشيئا عن المحور العمودي ويأخذ شكلا جرسيا كلما زادت قيمة v. ونبرهن أنه عند v كبير $(v \ge 30)$ فإنها تتبع التوزيع الطبيعي المعياري في الجداول الاحصائية، تعين نقطة (قيمة المتغيرة) ك 2 على المحور الأفقي (أنظر الرسم المقابل) من خلال v
بالإضافة إلى المساحة p على يسار ك 2 تحت المنحنى $p = P(X \le \chi^2 v;p))$. وأحيانا تحدد النقطة ك 2 بدلالة المساحة على يمينها $(\alpha = 1- p)$ لذلك نجد في كتب الاحصاء كل من الكتابتين: $\chi^2 \alpha,v$ و $\chi^2 p,v$

[3]3 

ج. توزيع ستيودنت (Student's t-distribution)

لتكن المتغيرتان العشوائيتان المستقلتان Y و Z حيث $Y\sim N(0, 1)$ و $Z \sim \chi v^2$ ؛ المتغيرة تتبع توزيع لها دالة الكثافة التالية:

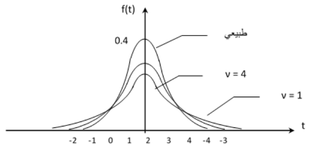
$$f(t)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)} \qquad -\infty<t<\infty$$

$\alpha > 0$ $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ / و نقول أن المتغيرة X تتبع توزيع ستيودنت ب vدرجة حرية ونكتب:

T ~ tv

خصائص توزيع ستيودنت 7[7]

E(T) = 0, V(T) = v/(v-2) si (v > 2)



نلاحظ أن منحنى tمتماثل حول المتوسط 0 مما يعني أن لكل نقطة موجبة t نقطة مناظرة لها سالبة حيث المساحة تحت المنحنى على يمين t تساوي المساحة تحت المنحنى علي يسار (t–)، ونكتب - = t1-p tp

بالإضافة إلى ذلك فإن منحنى(f(t يقترب من المنحنى الطبيعي المعياري كلما زادت قيمة v . وعموما، يعتبر الإحصائيون أن المنحنيان يتطابقان تقريبا عند 30≤ v.

في الجداول الاحصائية، تعين نقطة (قيمة المتغيرة) t من خلال vوالمساحة p على يسار t تحت المنحنى ((P(T ≤ tv;p = p. وأحيانا تحدد النقطة t بدلالة المساحة على يمينها (α = 1 - p) ونكتب : tp,v أو tα,v .5[5]

ج. تمرين

[17 ص 2 حل رقم]

أجب بصحيح أو خطأ

يعتبر التوزيع الطبيعي من التوزيعات المتناظرة

- ○ صحيح
- ○ خطأ

تمرين

تتمثل معالم التوزيع الطبيعي في:

- ☐ الوسيط والانحراف المعياري
- ☐ الوسط الحسابي والانحراف المعياري

تمرين

ماذا تمثل α

- ☐ مستوى الثقة
- ☐ مستوى المعنوية

أجب بصحيح أو خطأ

يقترب توزيع ستيودنت من التوزيع الطبيعي كلما زادت درجات الحرية.

- صحيح
- خطأ

باستعمال الجداول الاحصائية للتوزيع الطبيعي أوجد

$P(Z \leq z)$ حيث $z = 1, 2, 3$

خاتمة

إن عملية التوصيف الدقيق للتجارب العشوائية أمر بالغ الأهمية لا سيما عند التعامل مع ظروف متماثلة، حيث تم وضع قوانين التوزيع الاحتمالي حتى تسرع عملية الوصول إلى نتائج، فالتوزيعات الاحتمالية الخاصة بالمتغيرات المستمرة فالشائعة في الدراسات الاقتصادية تعطي رؤية أفضل للدراسة الظواهر الاقتصادية وتساعد متخذ القرار في عملية التنبؤ واتخاذ القرار

حل التمارين

> 1 (ص 10)

تمرين

الاحتمال هو 0.4 أو 40%.

تمرين

صحيح

تمرين

- [x] 0.1 عميل لكل دقيقة
- [ ] 10 عميل لكل دقيقة

أجب بصحيح أو خطأ

صحيح

> 2 (ص 13)

أجب بصحيح أو خطأ

- (●) صحيح
- ( ) خطأ

تمرين

- [ ] الوسيط والانحراف المعياري
- [x] الوسط الحسابي والانحراف المعياري

تمرين

- [ ] مستوى الثقة
- [x] مستوى المعنوية

أجب بصحيح أو خطأ

- (•) صحيح
- ( ) خطأ

باستعمال الجداول الاحصائية للتوزيع الطبيعي أوجد

0.8413 ،0.9772 ، 0.9986