مبادئ نظرية التوزيعات الاحتمالية

(212 احص)

Probability Distributions Theory

(STAT 212)

د. وليد المطيري

المتغير العشوائي (Random Variable):

لنفرض أن S هو فضاء العينة لتجربة عشوائية. وبالتالي فان المتغير العشوائي X هو دالة حقيقية معرفة على فضاء العينة.

ملاحظات:

- المتغير العشوائي X يعطي قيمة حقيقية وحيدة لكل عنصر من عناصر فضاء العينة S.
- المتغير العشوائي X هو تطبيق مجاله فضاء العينة S ومجاله المقابل هو مجموعة الأعداد الحقيقية R. $X: S \to R$
- اذا كانت $\omega \epsilon S$ نقطة عينة فان صورة ω تحت تأثير المتغير العشوائي X هي قيمة حقيقية أي أن: $X(\omega) \in R$.
- المجموعة $\{X \in R: X(\omega) = X, \omega \in R\}$ $X(S) =$ هي مدى التطبيق X وتسمى مجموعة القيم الممكنة للمتغير العشوائي X وكذلك $X(S) \subseteq R$

أنواع المتغيرات العشوائية:

1- المتغيرات العشوائية المتقطعة (المنفصلة)

2- المتغيرات العشوائية المتصلة (المستمرة)

1- المتغير العشوائي المتقطع:

يكون المتغير العشوائي X متغيرا عشوائيا متقطعا اذا كانت مجموعة القيم الممكنة له X(S) هي مجموعة متقطعه أو قابلة للعد. ويمكن أن تأخذ احدى الحالتين:

$$X(S) = \{x_1, x_2, \dots, x_n\}$$

$$X(S) = \{x_1, x_2, \dots\}$$

بشكل عام يرمز للمتغير العشوائي بالاحرف الانجليزية الكبيرة ويرمز للقيم التي يأخذها المتغير بالأحرف الانجليزية الصغيرة.

<u>أمثلة على المتغيرات العشوائية المتقطعة:</u>

.1

.2

.3

جدول التوزيع الاحتمالي:

- هو جدول مؤلف من سطرين يسهل عرض القيم الممكنة للمتغير العشوائي المتقطع X مع احتمالياتها المقابلة لها.

مثال: كون جدول توزيع احتمالي لـ X والذي يمثل عدد مرات اختيار رجل (male) في عينة عشوائية من ثلاثة أشخاص بفضاء عينه هو:

$$S = \{FFF, FFM, FMF, MFF, MMM, MFM, MMF, FMM\}$$

قانون التوزيع الاحتمالي للمتغير العشوائي المتقطع:

هو علاقة رياضية تربط بين قيم المتغير العشوائي X المتقطع والاحتمالات المقابلة لها. وكل قانون احتمالي يمكن تحويله الى جدول توزيع احتمالي.

يسمى أيضا القانون بدالة الكتلة الاحتمالية للمتغير العشوائي X ويرمز لها بالرمز f(x) وتعرف كالتالي:

$$f_X(x) = \begin{cases} p(X = x) & ; \ x \in X(S) \\ 0 & ; \ x \notin X(S) \end{cases}$$

<u>خواصها:</u>

1. $0 \leq f_X(x) \leq 1$
2. $\sum_{\forall x} f_X(x) = 1$

مثال

لتكن التجربة العشوائية هي رمي قطعة نقدية مرتين متتاليتين بشكل مستقل ولنفرض أن هذه العملية غير متزنه. بمعنى أن $p(H) = 1/3$ احتمالية ظهور صورة و $p(T) = 2/3$ احتمالية ظهور كتابة

ولنعرف المتغير العشوائي X على أنه عدد الصور الظاهرة في كلتا الرميتين. ,عليه أوجد:

1) فضاء العينة

2) دالة الكتلة الاحتمالية (تحقق من الشروط السابقة).

الحل:

مثال

لدينا القانون التالي للمتغير العشوائي المتقطع

$$p(X = x) = \frac{x+1}{21}, \qquad x = 0, 1, 2, 3, 4, 5$$

وضح اذا كان القانون السابق يحقق جميع شروط قانون التوزيع الاحتمالي؟

الحل:

تمرين:

- في تجربة رمي قطعتين من حجر النرد معا، وليكن X يمثل ما يلي:

1. X مجموع الوجهين الظاهرين.
2. X يمثل القيمة المطلقة للفرق بين الوجهين الظاهرين.

كون جدول التوزيع الاحتمالي في كلتا الحالتين.

التوقع (المتوسط الحسابي) للمتغير العشوائي المتقطع:

يرمز له بالرمز $E(X)$ أو μ_X ويعرف بالصيغة التالية:

$$\mu_X = E(X) = \sum_{x \in \mathrm{X}(S)} xf(x)$$

$$= \sum_{x \in \mathrm{X}(S)} xp(X = x)$$

$$= x_1 p(X = x_1) + x_2 p(X = x_2) + \ \ldots$$

- **<u>مثال:</u>**

أوجد التوقع أو المتوسط الحسابي للمثال السابق؟

بعض خواص التوقع:

- $E(a) = a$
- $E(X \pm a) = E(X) \pm a$
- $E(aX \pm b) = aE(X) \pm b$
- $E\big(g(X)\big) = \sum_{x \in X(S)} g(x) f(x)$

التباين للمتغيرات العشوائية المتقطعة:

اذا كان X متغيرا عشوائيا متقطعا توقعه (متوسطه الحسابي) μ_X فان تباينه يرمز له بالرمز σ_X^2 أو $var(X)$

$$\sigma_X^2 = var(X) = E(X-\mu_X)^2 = E(X^2) - \mu_X^2 = E(X^2) - [E(X)]^2$$

<u>أثبت أن :</u>

$$E(X-\mu_X)^2 = E(X^2) - [E(X)]^2$$

التباين للمتغيرات العشوائية المتقطعة:

- الانحراف المعياري هو الجذر التربيعي للتباين ويرمز له بالرمز σ_X.

- التباين هو نتيجة مباشرة من خواص التوقع وذلك لأن:

$$\sigma_X^2 = \sum_{x \in R_X} (x - \mu_X)^2 f_X(x)$$

مثال: اذا كانت لدينا دالة الكتلة الاحتمالية التالية:

x	0	1	2
P(X=x)	0.6	0.3	0.1

فأوجد المتوسط والتباين والانحراف المعياري؟

بعض خواص التباين:

بفرض أن X متغير عشوائي متقطع و a,b ثوابت فان:

- $var(a) = 0 \qquad (؟)$
- $var(x \pm a) = var(x)$
- $var(ax) = a^2 var(x)$
- $var(ax \pm b) = a^2 var(x)$

مثال

لدينا القانون التالي للمتغير العشوائي المتقطع

$$p(X=x)=\frac{x+1}{21}, \qquad x=0,1,2,3,4,5$$

فاذا كان $Y=2X+5$ ، فأوجد $E(Y)$ و $var(Y)$

الحل:

دالة التوزيع الاحتمالي التجميعية للمتغيرات العشوائية المتقطعة:

- هي دالة تعبر عن احتمالية ان يأخذ المتغير العشوائي X قيمة أقل أو تساوي قيمة ما من القيم الممكنة من قيم المتغير العشوائي.
- يرمز لها بالرمز F(x) وتعرف بالشكل التالي:

$$F(x) = p(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

<u>وتحقق الشروط التالية:</u>

1. $0 \leq F(x) \leq 1$ لأنها دالة احتمالية.
2. دالة غير متناقصة.

<u>ملاحظة</u>: اذا كانت F(x) هي دالة التوزيع الاحتمالي للمتغير العشوائي المتقطع X فان:

$$p(X > x) = 1 - p(X \leq x) = 1 - F(x)$$

- مثال: لتكن لدينا دالة الكتلة الاحتمالية للمتغير العشوائي المتقطع X كالاتي:

$$f(x) = \begin{cases} \frac{1}{55}x & ; \ x = 1, 2, 3, \dots, 10 \\ 0 & ; \quad O.W. \end{cases}$$

فأوجد دالة التوزيع التجميعية؟ ثم احسب $p(x \leq 3)$ $p(x > 2)$ $p(2 \leq x \leq 4)$

للمساعدة: من قوانين المتسلسلات $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$

الحل:

- مثال: لتكن لدينا دالة الكتلة الاحتمالية للمتغير العشوائي المتقطع X كالاتي:

$$f(x) = \begin{cases} \frac{1}{15}x & ; x = 1, 2, 3, 4, 5 \\ 0 & ; \quad O.W. \end{cases}$$

فأوجد جدول التوزيع التجميعي؟ ثم احسب $p(x \leq 2)$ $F(3)$

الحل:

x	1	2	3	4	5
$p(X = x)$					

x	1	2	3	4	5
$p(X \leq x)$					

حالات خاصة في حساب الاحتمال:

- اذا كان X هو متغير عشوائي متقطع عندئذ:

1. $p(X < x) = p(X \leq x) - p(X = x)$
2. $p(X \geq x) = 1 - p(X < x)$
3. $p(X > x) = 1 - p(X \leq x)$

تمرين:

- اذا كان لدينا جدول التوزيع الاحتمالي التالي:

x	4	5	6	7
P(X=x)	5/30	1/30	16/30	8/30

- احسب التالي:

1. $p(X < 5)$
2. $p(X > 4)$
3. $p(X \geq 5)$

العزوم للمتغيرات العشوائية المتقطعة:

إن العزم من المرتبة r للمتغير العشوائي يرمز له بالرمز α_r ويعطى بالعلاقة التالية:

$$\alpha_r = E(x^r) = \sum x^r p(X = x), \qquad r = 1, 2, 3, \dots, n$$

مثال:

- $\alpha_5 = E(x^5)$ عزم مرتبة خامسة
- $\alpha_3 = E(x^3)$ عزم مرتبة ثالثة
- $\alpha_2 = E(x^2)$ عزم مرتبة ثانية

نتيجة:

التوقع الرياضي (المتوسط الحسابي) هو العزم من المرتبة الأولى.

مثال: لدينا القانون الاحتمالي التالي للمتغير العشوائي المتقطع X:

$$p(X=x)=\frac{x+1}{21},\qquad x=0,1,2,3,4,5$$

احسب العزوم من المرتبة الثانية (α_2) والثالثة (α_3) والخامسة (α_5)؟

الحل:

نتيجة: التباين يمكن التعبير عنه عن طريق حساب العزوم من المرتبة الأولى والثانية.

تمرين: أثبت أن $var(x) = \sigma^2 = \alpha_2 - \alpha_1^2$ من أجل أي متغير عشوائي X.

الدوال المولدة للعزوم للمتغيرات العشوائية المتقطعة:

نرمز لهذه الدالة بـ $M_x(t)$ والتي من خلالها نعرف عدد من من العزوم إلى X.

<u>تعريف:</u>

ليكن X متغيرا عشوائيا متقطعا له دالة التوزيع الاحتمالي $f(x) = p(X = x)$ في الفضاء R_X فان الدالة المولدة للعزم $M_x(t)$ تعرف بالشكل التالي:

$$M_X(t) = \sum_{x \in R_X} e^{tx} f(x)$$

ومن العلاقة السابقة نلاحظ أن الدالة $M_x(t)$ تمثل القيمة المتوقعة بالنسبة إلى e^{tx}

$$M_X(t) = E(e^{tx})$$

عندما نتعامل مع الدالة المولدة للعزوم نفترض أنها موجودة لجميع قيم t بحيث أن $-h < t < h$ لبعض h

نظرية:

وجود الدالة المولدة للعزوم $M_x(t)$ يؤدي الى أن مشتقاتها بالنسبة إلى t لجميع الرتب عندما t=0 تكون موجودة. أي أن:

$$M'_X(t) = \frac{d}{dt}\{M_X(t)\} = \sum_{x \in R_X} x e^{tx} f(x)$$

$$M''_X(t) = \frac{d^2}{dt^2}\{M_X(t)\} = \sum_{x \in R_X} x^2 e^{tx} f(x)$$

$$M^r_X(t) = \frac{d^r}{dt^r}\{M_X(t)\} = \sum_{x \in R_X} x^r e^{tx} f(x)$$

وعندما نضع t=0 نحصل على:

$$M'_X(0) = \sum_{x \in R_X} x e^{0x} f(x) = \sum_{x \in R_X} x f(x) = E(X) = \alpha_1$$

$$M''_X(0) = \sum_{x \in R_X} x^2 e^{0x} f(x) = \sum_{x \in R_X} x^2 f(x) = E(X^2) = \alpha_2$$

مثال: لتكن الدالة المولدة للعزوم للمتغير العشوائي X معرفة كالتالي:

$$M_x(t) = \frac{1}{6} + \frac{3}{6}e^{t} + \frac{2}{6}e^{3t}$$

فأوجد المتوسط الحسابي والتباين؟

تمرين: لتكن الدالة المولدة للعزوم للمتغير العشوائي المتقطع X معرفة كالتالي:

$$M_x(t) = \frac{1}{10}\left(e^{t} + 2e^{3t} + 4e^{5t} + 2e^{7t} + e^{9t}\right)$$

فأوجد:

1. قيم X
2. جدول التوزيع الاحتمالي
3. المتوسط الحسابي
4. التباين

التوزيعات الاحتمالية المتقطعة

محاولة برنولي:

- هي تجربة عشوائية لها نتيجتين فقط مثلاً نجاح (S) أو فشل (F).
- فضاء عينة تجربة برنولي $S = \{S,F\}$.
- احتمال النجاح يرمز له بالرمز $p(S) = p$
- احتمال الفشل يرمز له بالرمز $p(F) = q=1-p$
- أمثلة على محاولات برنولي:

1.
2.
3.

توزيع برنولي (Bernoulli Distribution):

اذا رمزنا للنجاح بـ 1 والفشل بـ 0 فبالتالي فضاء العينة يصبح S={0,1} و p(X=1) = p و p(X=0)=q=1-p

فان دالة الكتلة الاحتمالية للمتغير العشوائي X الذي يتبع توزيع برنولي هي:

$$f(x) = p(X = x) = \begin{cases} p^x(1-p)^{1-x} & ; \quad x = 0,1 \\ 0 & ; \quad x \neq 0,1 \end{cases}$$

أو بمعنى آخر:

$$f(x) = p(X = x) = \begin{cases} p & ; \quad x = 1 \\ 1-p & ; \quad x = 0 \\ 0 & ; \quad x \neq 0,1 \end{cases}$$

توزيع ذي الحدين (Binomial Distribution):

لنفرض أن التجربة العشوائية تتكون من تكرار محاولة برنولي عدد n من المرات وتحقق الشروط التالية:

- $n > 1$
- **المحاولات مستقلة**
- **احتمال النجاح $p(S) = p$ ثابت لجميع المحاولات.**

فعليه نعرف المتغير العشوائي X على أنه عدد مرات النجاح عند تكرار تجربة برنولي، وبالتالي تكون مجموعة القيم الممكنة :

$$X(S) = \{0, 1, 2, 3, \dots, n\}$$

ودالة الكتلة الاحتمالية للمتغير العشوائي X هي:

$$f(x) = p(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}; & x = 0,1,2,\dots,n \\ 0 \ ; & O.W. \end{cases}$$

ويسمى هذا التوزيع بتوزيع ذي الحدين بالمعالم n, p وتكتب كالتالي:

$$X \sim Bin(n, p)$$

توزيع ذي الحدين (Binomial Distribution):

ملاحظات:

- توزيع برنولي هو حالة خاصة من توزيع ذي الحدين عندما n=1.
- اذا كان X المتغير العشوائي لعدد مرات النجاح و Y المتغير العشوائي لعدد مرات الفشل وكان

$$X \sim Bin(n, p)$$

فان Y يتبع توزيع ذي الحدين كالتالي

$$Y \sim Bin(n, 1-p)$$

التوقع والتباين لتوزيع ذي الحدين:

اذا كان $X \sim Bin(n,p)$ فان:

$$\mu_X = E(X) = n.p$$

$$\sigma_X^2 = var(X) = n.p.(1-p)$$

الدالة المولدة للعزوم لتوزيع ذى الحدين:

$$M_X(t) = (1 - p + pe^t)^n$$

<u>مثال:</u>

في تجربة عشوائية تم رمي ست زهرات نرد، ولنفرض أننا نراقب ظهور الوجه الذي يحمل الرقم 3 أوجد دالة الكتلة الاحتمالية الخاصة بـ X حيث أن :

X: ظهور الوجه الذي يحمل الرقم 3

تكملة للمثال السابق: احسب الاحتمالات التالية:

- احتمال ظهور الوجه 3 مره واحدة.
- احتمال ظهور الوجه 3 على الأقل 3 مرات.

توزيع بواسون (Poisson Distribution):

التجارب التي تعطينا عدد النجاحات في فترة زمنية معينة أو منطقة محددة تسمى تجارب بواسون.
الفترة الزمنية يمكن أن تكون مثلا: ثانية أو دقيقة أو يوما أو أسبوعا أو شهرا أو غير ذلك.
المنطقة المحددة يمكن أن تكون مثلا: كصفحة من كتاب أو كمتر مربعا من مساحة.
أمثلة على تجارب بواسون:

- عدد الزبائن الذين يدخلون الى مكتب البريد كل خمس دقائق.
- عدد حوادث السيارات على طريق ما في أسبوع.
- عدد المكالمات الهاتفية التي تصل الى مكتب ما كل عشر دقائق.

توزيع بواسون (Poisson Distribution):

اذا كان X متغير عشوائي يمثل عدد النجاحات في فترة زمنية معينة أو منطقة محددة، فيمكننا القول أنه يتبع توزيع بواسون ودالة توزيعه الاحتمالية كالتالي:

$$f(x) = p(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \; ; \quad x = 0, 1, 2, 3, \dots$$

حيث λ تمثل معدل عدد النجاحات في الفترة الزمنية المعينة أو المنطقة المحددة.

ويكتب X يتبع توزيع بواسون بمعلومية المعلمة λ بالشكل التالي:

$$X \sim Pois(\lambda)$$

التوقع (المتوسط الحسابى) التباين لتوزيع بواسون:

$$\mu_X = E(X) = \lambda$$
$$\sigma_X^2 = var(X) = \lambda$$

الدالة المولدة للعزوم لتوزيع بواسون:

$$M_X(t) = e^{\lambda(e^{t}-1)}$$

مثال:

اذا كان عدد المكالمات الهاتفية القادمة الى محطة هاتف هو متغير عشوائي X وكان معدل عدد المكالمات خلال ساعه يساوي 3 مكالمات، فاحسب التالي:

- احتمال قدوم مكالمتين فقط الى المحطة.
- احتمال قدوم مكالمتين على الأقل.

مثال:

اذا كان معدل عدد حوادث السيارات عند إشارة ضوئية معينة في أسبوع هو 3 حوادث، فاحسب التالي:

- احتمال عدم حدوث أي حادث عند تلك الإشارة في أسبوع معين.
- احتمال حدوث حادثين على الأكثر في أسبوع معين.

2- المتغير العشوائي المتصل:

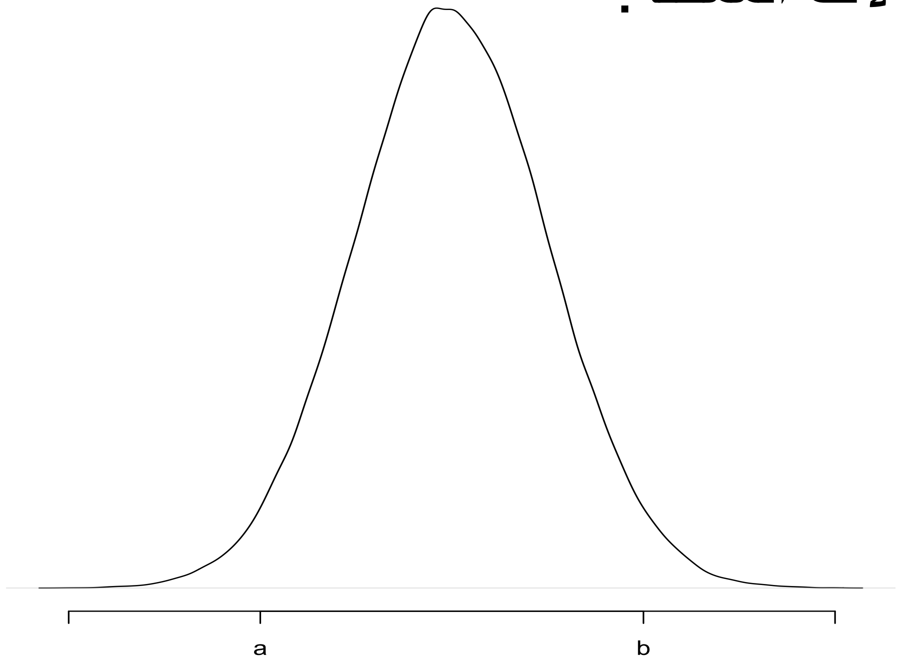
- يمكن أن يعرف المتغير العشوائي X بأنه متغير عشوائي متصل اذا كانت جميع القيم الممكنة له هي عبارة عن قيم لأعداد حقيقية من فترة أو اتحاد فترات.

أمثلة:

1.
2.
3.

دالة الكثافة الاحتمالية للمتغير العشوائي المتصل:

لأي متغير عشوائي متصل X يوجد دالة حقيقية موجبة $f_X(x)$ تسمى دالة الكثافة الاحتمالية التي من خلالها نستطيع إيجاد الاحتمالات الممكنة.



المساحة تحت منحنى هذه الدالة تعطي احتمال وقوع المتغير العشوائي X لفترة معينة:

$$p(a < X < b) = \int_a^b f_X(x)dx$$

خصائص دالة كثافة الاحتمال:

- الدالة f(x) موجبة داخل الفترة (a, b) أي أن $x \in (a,b); f(x) \geq 0$
- التكامل على حدود فترة المتغير من الحد الأدنى a الى الحد الأعلى b يعبر عن مجموع الاحتمالات الكلية والتي لابد أن تساوي الواحد. أي أن:

$$\int_a^b f(x)dx = 1$$

- احتمالية الحصول على قيمة معينة محددة في المتغيرات العشوائية المتصلة تساوي الصفر

اذا كان $x \in (a,b)$ فان $p(X = x) = 0$

أثبت ذلك؟

مثال: أثبت أن الدالة أدناه هي دالة كثافة احتمالية:

$$f(x) = \begin{cases} 2x & ; \quad 0 < x < 1 \\ 0 & ; \quad O.W. \end{cases}$$

مثال: اذا كان لدينا الدالة الكثافة الاحتمالية أدناه فاحسب قيمة c التي تجعل منها دالة كثافة احتمالية:

$$f(x) = \begin{cases} 2x & ; \quad 0 < x < 1 \\ 0 & ; \quad O.W. \end{cases}$$

<u>مثال:</u> اذا كان لدينا الدالة الكثافة الاحتمالية أدناه فاحسب قيمة c التي تجعل منها دالة كثافة احتمالية:

$$f(x)=\begin{cases} cx(10-x); & 0<x<10 \\ 0 \quad ; & O.W. \end{cases}$$

ثم احسب $p(X<3)$

المتوسط الحسابي والتباين للمتغير العشوائي المتصل:

اذا كانت $f_X(x)$ دالة الكثافة الاحتمالية للمتغير العشوائي المتصل X بحيث أن $x \in (a,b)$ فان المتوسط الحسابي (التوقع) للدالة h(x) تأخذ الصيغة التالية:

$$E(h(x)) = \int_a^b h(x)f(x)\,dx$$

وبالتالي يمكننا كتابة صيغة المتوسط الحسابي والتباين كالآتي:

$$\mu = E(X) = \int_a^b xf(x)\,dx$$

$$E(X^2) = \int_a^b x^2f(x)\,dx$$

$$var(X) = \sigma^2 = E(X^2) - \big(E(X)\big)^2$$

مثال: اذا كانت لدينا دالة الكثافة الاحتمالية التالية:

$$f(x) = \begin{cases} 0.006\,x(10-x); & 0 < x < 10 \\ 0 & ; O.W. \end{cases}$$

فاحسب المتوسط الحسابي والتباين؟

<u>مثال:</u> اذا لدينا الدالة التالية:

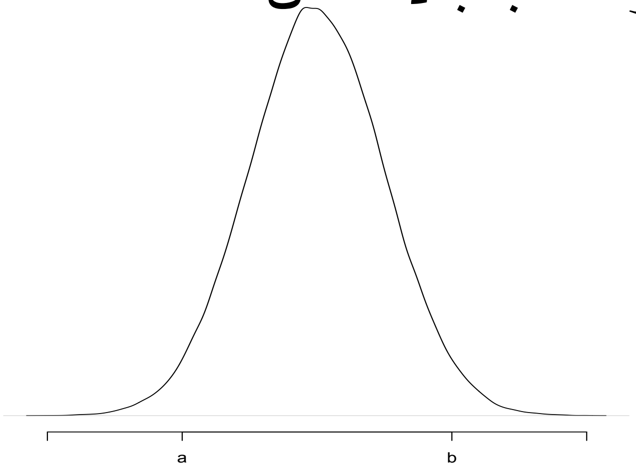
$$f(x)=\begin{cases}\frac{1}{c}(x-2)(x+3); 3<x<7\\ 0 \qquad ; \quad O.W.\end{cases}$$

فأوجد التالي:

- قيمة c التي تجعل f(x) دالة كثافة احتمالية
- أوجد المتوسط الحسابي والتباين
- احسب الاحتمالات التالية:

- $P(3<X<7)$
- $P(4<X<5)$
- $P(X<6)$
- $P(X>3)$
- $P(X=3)$
- $P(X=7)$

دالة التوزيع التجميعية للمتغير العشوائي المتصل:



- نرمز لهذه الدالة بـ F(x) وتسمى دالة التوزيع التجميعية للمتغير العشوائي X وتحسب بالاحتمال التالي:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

خواصها:

- $\lim_{x\to\infty} F(x) = 1, \quad \lim_{x\to-\infty} F(x) = 0$
- دالة متزايدة بالنسبة للمتغير X أي لكل $x_1 < x_2$ فان $F(x_1) < F(x_2)$
- اذا كان $x_1 < x_2$ فان

$$p(x_1 < X < x_2) = F(x_2) - F(x_1)$$

- اذا كانت F(x) دالة التوزيع التجميعية للمتغير العشوائي X فان :

$$p(X > x) = 1 - p(X \le x) = 1 - F(x)$$

مثال: اذا كانت لدينا دالة الكثافة الاحتمالية التالية:

$$f(x) = \begin{cases} 0.006x(10 - x) & ; \quad 0 < x < 10 \\ 0 & ; \quad O.W. \end{cases}$$

فأوجد دالة التوزيع التجميعية F(x)؟

مثال: أوجد دالة التوزيع التجميعية لدالة الكثافة الاحتمالية التالية:

$$f(x) = \begin{cases} \dfrac{1}{10}(x+3) & ; \quad 1 < x < 3 \\ 0 & ; \quad O.W. \end{cases}$$

التوزيعات الاحتمالية المتصلة

التوزيع المنتظم (Uniform Distribution):

- توزيع له دالة احتمال ثابته.
- يستخدم في حالة الظواهر التي يمكن أن تحدث بشكل مستمر.
- اذا كان X متغيراً عشوائياً يتبع التوزيع المنتظم للفترة [a,b] فان دالة كثافته الاحتمالية هي:

$$f(x) = \{\frac{1}{b-a}; \quad a < x < b$$

ويكتب: $x \sim U(a,b)$

معالم هذا التوزيع هما المعلمتان المحددتان لفترة قيم X الممكنة a و b.

التوزيع المنتظم (Uniform Distribution):

- المتوسط الحسابي:

$$\mu = E(x) = \frac{a+b}{2}$$ أثبت ذلك؟

- التباين:

$$\sigma^2 = var(x) = \frac{(b-a)^2}{12}$$ أثبت ذلك؟

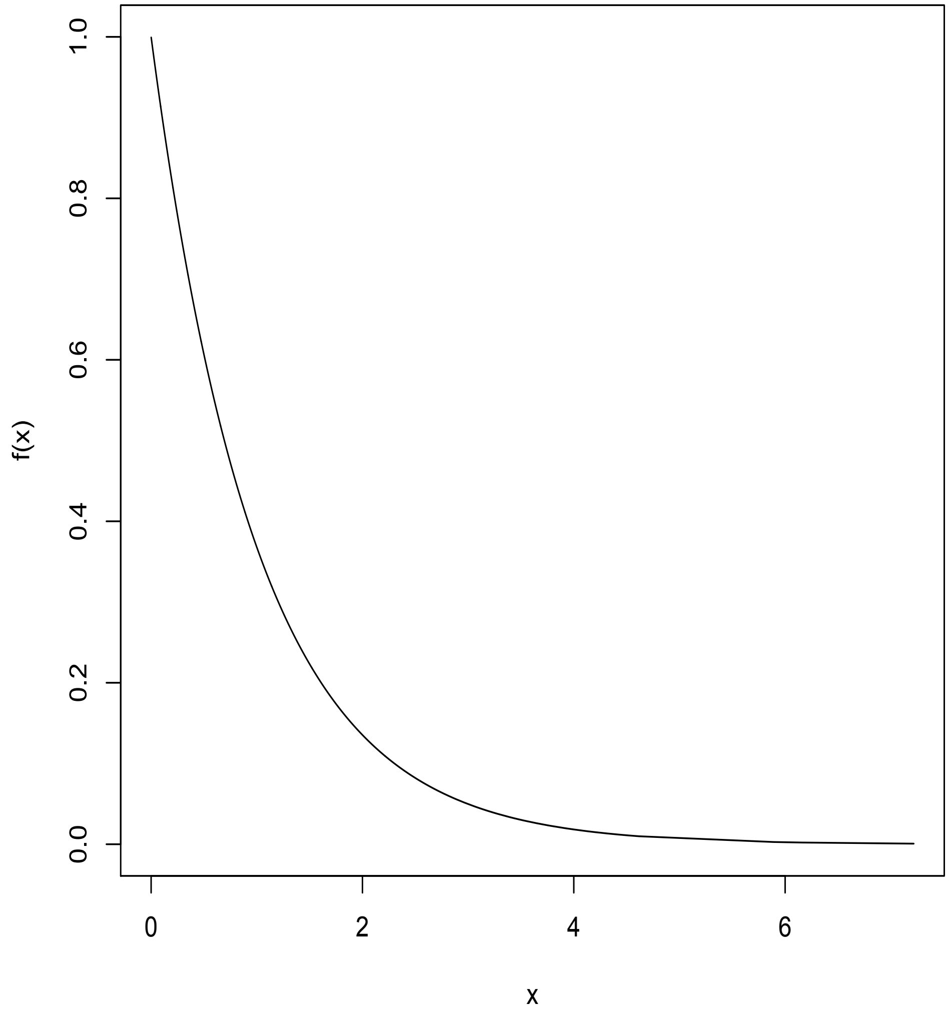
- دالة التوزيع التجميعية:

$$F(x) = \frac{x-a}{b-a}$$ أثبت ذلك؟

التوزيع الأسي (Exponential Distribution):

- توزيع احتمالي مستمر اشتق اسمه من الدالة الأسية.
- عادة ما يستخدم التوزيع الأسي في مسائل متعلقة بقياس الزمن. من ذلك مدة خدمة شباك البريد، مدة المكالمة هاتفية، مدة تفريغ باخرة الشحن، مدة تصليح آلة، مدة انتظار زبون قبل الحصول على الخدمة.
- للتوزيع الأسي علاقة بتوزيع بواسون، فإذا كان وقوع أحداث ما يتبع توزيع بواسون، 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 "ب" تتبع التوزيع الأسي.

التوزيع الأسي (Exponential Distribution):



اذا كان المتغير العشوائي X يتبع التوزيع الأسي الذي مداه هو $\infty < x < 0$ فان دالة كثافته الاحتمالية هي:

$$f(x) = \theta e^{-\theta x}; \quad 0 < x < \infty \quad and\ \theta > 0$$

ويكتب: $x \sim Exp(\theta)$

هذا التوزيع لديه معلمة واحدة فقط وهي θ.

التوزيع الأسي (Exponential Distribution):

المتوسط الحسابي:

$$\mu = E(x) = \frac{1}{\theta}$$

التباين:

$$\sigma^2 = var(x) = \frac{1}{\theta^2}$$

دالة التوزيع التجميعية:

$$F(x) = p(X \leq x) = \int_0^x f(t)dt \ = 1 - e^{-\theta x}$$

$$\mu = E(x) = \int_0^\infty x f(x) dx = \int_0^\infty x\theta e^{-\theta x} dx = \theta \int_0^\infty x e^{-\theta x} dx$$

باستخدام التكامل بالتجزئة:

$$u = x \qquad dv = e^{-\theta x} dx$$

$$du = dx \qquad v = \frac{-1}{\theta} e^{-\theta x}$$

يصبح:

$$\theta \int_0^\infty x e^{-\theta x} dx = \theta \left(\frac{-x}{\theta} e^{-\theta x} \Big|_0^\infty + \int_0^\infty \frac{1}{\theta} e^{-\theta x} dx \right)$$

$$= \left(-x e^{-\theta x} \Big|_0^\infty + \left[\left(\frac{-1}{\theta}\right) e^{-\theta x} \right] \Big|_0^\infty \right) = 0 + \left[0 - \frac{-1}{\theta} \right] = 1/\theta$$

تمرين:

أثبت قانون التباين وكذلك دالة التوزيع التجميعية للتوزيع الأسي؟

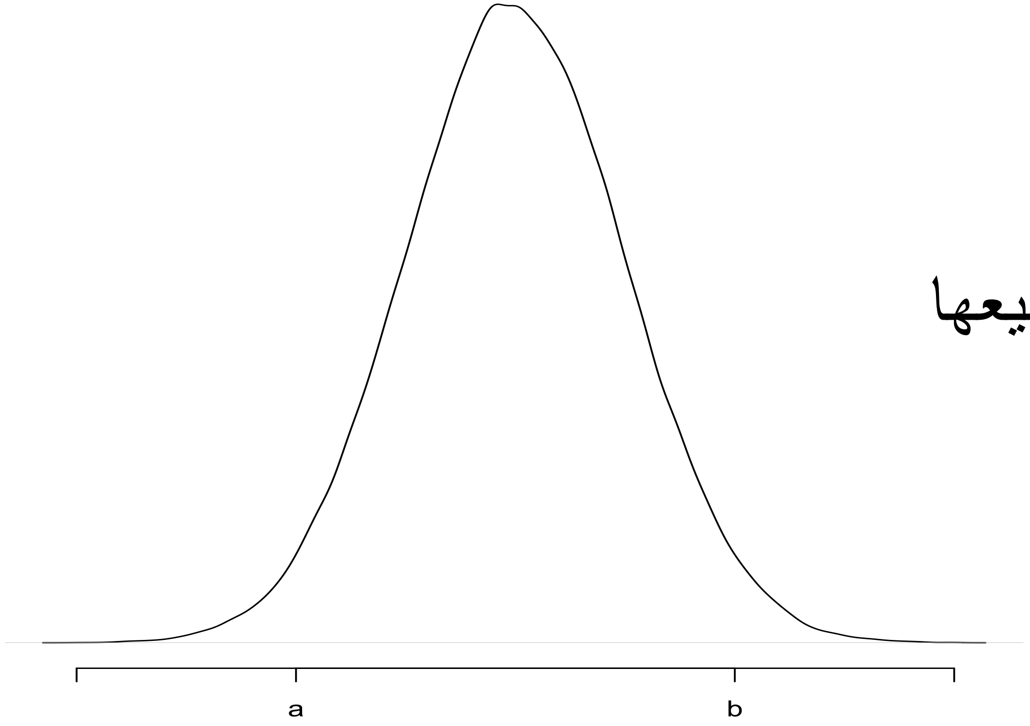
مثال:

اذا كانت الفترة الزمنية لإنهاء خدمة العميل بالبنك تتبع توزيع أسي بمتوسط دقيقتين فأوجد ما يلي:

1. **دالة الكثافة الاحتمالية المعبرة عن الفترة الزمنية لإنهاء خدمة العميل؟**
2. **ما هو احتمال انهاء خدمة العميل في أقل من 5 دقائق؟**
3. **ما هو احتمال انهاء خدمة العميل في مدة أكبر من 5 دقائق؟**

التوزيع الطبيعي (Normal Distribution):

- **التوزيع الطبيعي** هو توزيع احتمالي مستمر كثير الانتشار والاستعمال.
- يستخدم غالباً تقريباً أولياً لوصف المتغيرات العشوائية التي تميل إلى التمركز حول قيمة متوسطة وحيدة.
- دالة كثافته الاحتمالية تأخذ شكل الجرس.
- تكون متناظره حول المتوسط الحسابي.
- متوسطه الحسابي ووسيطه الحسابي ومنواله جميعها متساوية.



التوزيع الطبيعي (Normal Distribution):

- اذا كان المتغير العشوائي X يتبع التوزيع الطبيعي ومداه $-\infty < x < \infty$ فان دالة كثافته الاحتمالية هي:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} ; -\infty < x < \infty ; \sigma > 0$$

- ويكتب: $x \sim N(\mu, \sigma^2)$
- كلما زادت قيمة σ^2 كلما زاد تشتت البيانات، والعكس صحيح.
- التوزيع الطبيعي له معلمتان وهما: μ و σ^2 حيث:

$$\mu = E(x) \qquad \sigma^2 = var(x)$$

التوزيع الطبيعي المعياري (Standard Normal Distribution):

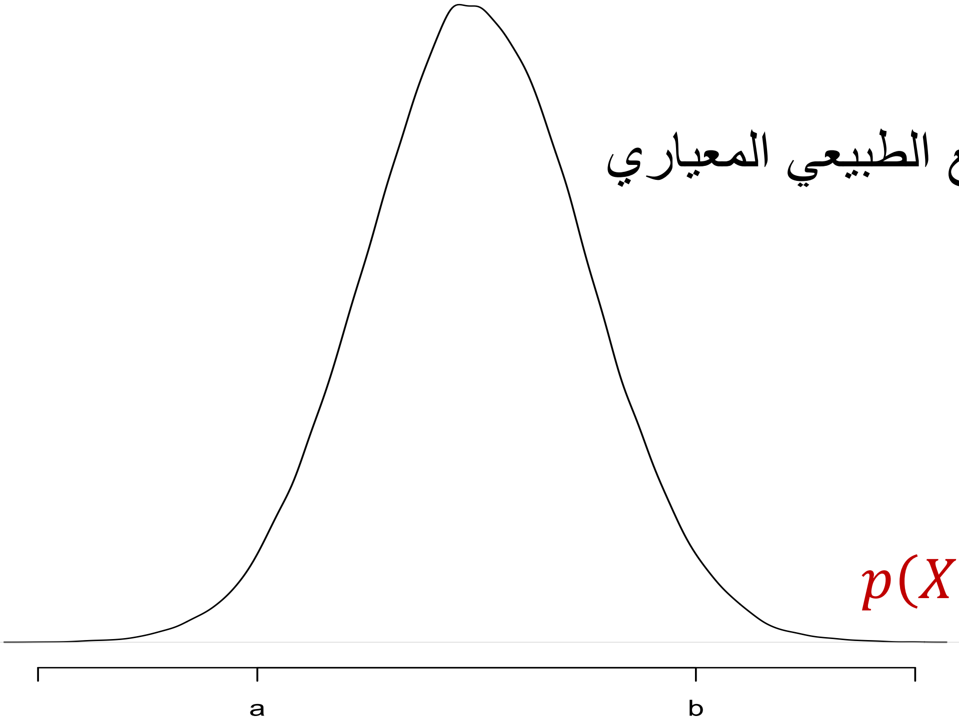
- اذا كان المتغير العشوائي Z يتبع التوزيع الطبيعي المعياري **فان متوسطه الحسابي لابد أن يساوي الصفر وتباينه يساوي الواحد** وتصبح دالة كثافته الاحتمالية كالتالي:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}; -\infty < z < \infty$$

- ويكتب: $z \sim N(\mu = 0, \sigma^2 = 1)$
- اذ كان $x \sim N(\mu, \sigma^2)$ فتتم معايرته الى $z \sim N(0,1)$ باستخدام القانون التالي:

$$z = \frac{x - \mu}{\sigma}$$

إيجاد الاحتمالات للتوزيع الطبيعي:



- مساحة ما تحت المنحنى اذا كان X يتبع التوزيع الطبيعي أو Z يتبع التوزيع الطبيعي المعياري تساوي الواحد.
- لحساب احتمالية أن تكون X أقل من قيمة معينة a تكتب كالتالي:

$$p(X \leq a) = \int_{-\infty}^{a} f(x)dx = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

- هذا التكامل يعبر عن المساحة المحصورة تحت المنحنى التي تكون فيها X أقل من قيمة a.
- ولصعوبة إيجاد حل للتكامل تتم معايرة التوزيع الطبيعي للتوزيع الطبيعي المعياري ومنها يمكننا إيجاد القيمة الاحتمالية من جداول التوزيع الطبيعي المعياري.

إيجاد الاحتمالات للتوزيع الطبيعي:

- تكتب: $p(X \leq a) = p\left(Z \leq \frac{a-\mu}{\sigma}\right) = \Phi(a)$
- $\Phi(z) = p(Z < z) = p(Z \leq z)$
- $p(Z > z) = 1 - p(Z \leq z) = 1 - \Phi(z)$
- $p(a < Z < b) = p(Z < b) - p(Z < a) = \Phi(b) - \Phi(a)$
- $p(Z = a) = 0$

جداول التوزيع الطبيعي المعياري

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

مثال: اذا كان المتغير العشوائي X يمثل الطول في أحد المجتمعات البشرية ويتبع التوزيع الطبيعي بمتوسط 165 سم وانحراف معياري 5 سم فأوجد التالي:

1. القيمة المعيارية للقيمة $X=172$.
2. قيمة X اذا كانت قيمتها المعيارية تساوي -0.52

مثال: لنفرض أن مستوى هيموجلوبين الدم في أحد المجتمعات البشرية يتبع التوزيع الطبيعي بمتوسط 16 وانحراف معياري 0.9:

- اذا اخترنا أحد الأشخاص بشكل عشوائي ما هو احتمال أن يكون مستوى الهيموجلوبين في الدم لديه أكبر من 14؟
- ماهي نسبة الأشخاص الذين مستوى هيموجلوبين الدم لديهم أكبر من 14؟
- ماهي نسبة الأشخاص الذين مستوى هيموجلوبين الدم لديهم تتراوح بين 14 و 18؟

التوزيعات الاحتمالية الثنائية

التوزيعات الثنائية:

- تطرقنا سابقاً الى التوزيعات ذات المتغير العشوائي الواحد حيث أن المشاهدات تأخذ متغير واحد فقط، وسوف نتطرق هنا الى حالات أعمق حيث المشاهدات تأخذ أو تتطلب متغيرين أو أكثر.
- هذه المتغيرات اما أن تكون كمية مستمرة أو نوعية متقطعة أو تكون مختلطة (متقطعة ومستمرة).
- مثال: في الدراسات الطبية:
- X: مشاهدات ضغط الدم، Y: معدل النبض.

المتغيرات العشوائية المتقطعة المشتركة:

- **مثال:**

حاوية تحتوي على خمس مواد:

2 سليمة (S).

2 معيبة عيب بسيط (M).

1 معيبة عيب رئيسي (F).

سحبت منها عينة عشوائية مكونة من مادتين بدون ارجاع، احسب فضاء العينة وأوجد القيم المرتبطة بالحدثين:

X: عدد المواد السليمة.

Y: عدد المواد المعيبة.

$$S = \{ \qquad \}$$

(x, y)					
$p(X = x, Y = y)$					

- لنفرض أن العينة المختارة في المثال السابق تتكون من ثلاث مواد، أكمل المطلوب كما في المثال السابق:

$$S = \{ \qquad \}$$

(x, y)					
$p(X = x, Y = y)$					

تعريف: ليكن (X,Y) متغير عشوائي حيث:

$$X = \left(x_1, x_2, x_3, \dots, x_{n_x}\right)$$
$$Y = \left(y_1, y_2, y_3, \dots, y_{n_y}\right)$$

ومعرف على الفضاء $R_{x,y}$ حيث:

$$R_{x,y} = \{\left(x_i, y_j\right); i = 1,2, \dots, n_x, j = 1,2, \dots, n_y\}$$

فان $p(X = x, Y = y)$ ويرمز لها بالرمز $f(x,y)$ تسمى دالة التوزيع الاحتمالي المشترك للمتغيرين العشوائيين X, Y.

وتحقق الخواص التالية:

- $0 \le f(x,y) \le 1$
- $\sum_{(x,y)\in R_{x,y}} \sum f(x,y) = 1$
- $p\left((x,y) \in A\right) = \sum_{(x,y)\in R_{x,y}} \sum f(x,y)$

تكملة للمثال السابق:

ما هو احتمال أن تكون العينة المسحوبة بها على الأقل مادة واحدة بعيب بسيط وعلى الأكثر مادة واحدة سليمة.

مثال:

اذا كانت دالة الاحتمال المشتركة بين عدد الأطفال في الأسرة التي بها ما بين طفل الى ثلاثة أطفال (X: x=1,2,3) وعدد الوحدات المستهلكة لحليب الأطفال الجاف من نوع معين كل أسبوع لكل أسرة (Y: y=0,1,2) تأخذ الصورة التالية:

$$f(x,y)=\frac{y-0.5x+2}{18}; \qquad x=1,2,3\ ,y=0,1,2$$

كون جدول التوزيع الاحتمالي المشترك؟

Y

X		0	1	2
	1			
	2			
	3			

دالة التوزيع التجميعية المشتركة للمتغيرات العشوائية المتقطعة:

اذا كان (X,Y) متغيرين عشوائيين متقطعين فان دالة التوزيع التجميعية الثنائية المشتركة لـ X وY هي:

$$F(x,y) = p(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(u,v)$$

وخواصها:

- $F(\infty,\infty) = 1$
- $F(-\infty,y) = F(x,-\infty) = 0$
- $F(x,y)$ دالة غير متناقصة لكل متغير متقطع.

مثال:

اذا كانت دالة التوزيع الاحتمالي المشتركة لـ X,Y هي:

$$f(x,y)=\begin{cases}\binom{2}{y}0.5^{x+2} & ; x=1,2,\dots,\infty;\ y=0,1,2\\ 0\ ; & O.W.\end{cases}$$

أوجد دالة التوزيع التجميعية المشتركة؟

المتغيرات العشوائية المتصلة المشتركة:

اذا فرضنا أن X,Y متغيرين عشوائيين متصلين فان f(x,y) تسمى دالة الكثافة الاحتمالية المشتركة للمتغيرين العشوائيين المتصلين X,Y اذا حققت الشروط التالية:

- $f(x,y) \geq 0 \;\forall\; x,y$
- $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dy\,dx = 1$

<u>مثال:</u> لدينا الدالة المشتركة التالية:

$$\boldsymbol{f(x,y) = \begin{cases} \frac{1}{3}(x+y); & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0; & O.W. \end{cases}}$$

هل هذه دالة كثافة احتمالية مشتركة منتظمة؟

مثال:

تحقق ما اذا كانت الدالة التالية هي دالة كثافة احتمالية منتظمة؟

$$f(x,y)=\begin{cases} xe^{-y}; & 0\le x\le 4,\ 0\le y\le 3 \\ 0; & O.W \end{cases}$$

مثال:

لدينا دالة الكثافة الاحتمالية التالية:

$$f(x,y)=\begin{cases}\dfrac{cy}{\sqrt{x}}; & 0\le x\le 4, 0\le y\le 3\\ 0; & O.W\end{cases}$$

أوجد قيمة c التي تجعل من الدالة السابقة دالة كثافة احتمالية منتظمة؟

دالة التوزيع التجميعية المشتركة للمتغيرات العشوائية المتصلة:

اذا كان X,Y متغيرين عشوائيين متصلين فان دالة التوزيع التجميعية الثنائية المشتركة لـ X وY هي:

$$F(x,y)=p(X\le x,Y\le y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\,du\,dv$$

وخواصها:

- $F(\infty,\infty)=1$
- $F(-\infty,y)=F(x,-\infty)=0$
- $F(x,y)$ دالة غير متناقصة لكل متغير متصل.

مثال:

اذا علمت أن:

$$f(x,y)=\begin{cases}4xy; & 0\le x\le 1, 0\le y\le 1\\ 0; & O.W\end{cases}$$

فأوجد التالي:

- F(x, y)
- $F(0.1,0.2)=p(X\le 0.1,Y\le 0.2)$

حالات خاصة فى حساب احتمال المتغيرات العشوائية المتصلة المشتركة:

- $p(X = x, Y = y) = 0$
- $p(X \le x, Y \le y) = p(X < x, Y \le y) = p(X \le x, Y < y) = p(X < x, Y < y)$

مثال:

اذا علمت أن:

$$f(x,y) = \begin{cases} \frac{1}{116}(y\sqrt{x} + 2x); & 1 \le x \le 5, 1 \le y \le 4 \\ 0; & O.W \end{cases}$$

فأوجد F(x, y)؟

وكذلك أوجد:

- $p(1 \leq X < 3,\ 1 \leq Y \leq 2)$
- $p(1 \leq X \leq 4,\ \ Y \leq 5)$
- $p(X \leq 1,\ \ 2 < Y < 3)$
- $p(2 < X < 4,\ \ 2 < Y < 3)$
- $p(0 \leq X < 5, 4 < Y < 6)$

التوقع الرياضي للمتغيرات العشوائية المشتركة:

تأخذ الصيغة $E\left(g(x,y)\right)$ حيث x ,y متغيرين عشوائيين و $g(x,y)$ هي دالة معرفة على المتغيرين العشوائيين x ,y.

مثال: اذا كان x ,y متغيرين عشوائيين فان الدالة $g(x,y)$ يمكن أن تأخذ:

$$g(x,y) = 2xy$$

$$g(x,y) = 2x + y$$

$$g(x,y) = x^2 + y^2$$

التوقع الرياضي للمتغيرات العشوائية المشتركة:

ويمكن تعريف التوقع الرياضي بشكل عام لأي دالة معرفة على المتغيرين العشوائيين x ,y كالتالي:

<u>فى حالة المتغيرات العشوائية المتقطعة الثنائية:</u>

$$E\big(g(x,y)\big) = \sum_x \sum_y g(x,y) f(x,y) = \sum_x \sum_y g(x,y) p(X = x, Y = y)$$

<u>فى حالة المتغيرات العشوائية المتصلة الثنائية:</u>

$$E\big(g(x,y)\big) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx dy$$

حيث $f(x,y)$ هي دالة الكثافة الاحتمالية المشتركة للمتغيرين العشوائيين x ,y.

مثال (على المتغيرات العشوائية المتقطعة الثنائية):

لنفرض أن لدينا جدول التوزيع الاحتمالي المشترك التالي:

أوجد التوقع الرياضي $E(g(x,y))$ اذا كانت:

- $g(x,y) = x + y$

		X	
		0	1
Y	0	1/7	2/7
	1	0	3/7
	2	1/7	0

مثال (على المتغيرات العشوائية المتقطعة الثنائية):

لنفرض أن لدينا جدول التوزيع الاحتمالي المشترك التالي:

أوجد التوقع الرياضي $E(g(x,y))$ اذا كانت:

- $g(x,y) = x + 2y$

		X	
		0	1
Y	0	1/7	2/7
	1	0	3/7
	2	1/7	0

تمرين:

أوجد $E(g(x,y))$ لجدول التوزيع الاحتمالي السابق اذا علمت أن:

- $g(x,y) = x^2 + e^y$
- $g(x,y) = (xy)^3$
- $g(x,y) = e^x e^y$
- $g(x,y) = e^x + e^y$

مثال (على المتغيرات العشوائية المتصلة الثنائية):

ليكن:

$$f(x,y)=\begin{cases}4xy;\, 0<x<1,\, 0<y<1\\ 0;\, O.W.\end{cases}$$

فاحسب $E(g(x,y))$ عندما تكون $g(x,y)=x+y^2$

مثال (على المتغيرات العشوائية المتصلة الثنائية):

ليكن:

$$f(x,y)=\begin{cases}4xy;0<x<1,0<y<1\\ 0;O.W.\end{cases}$$

فاحسب $E(g(x,y))$ عندما تكون $g(x,y)=\frac{x}{y}+\frac{y^2}{x^2}$

تمرين على المثال السابق:

احسب $E(g(x,y))$ عندما تكون:

- $g(x,y) = x^2y^3$
- $g(x,y) = \frac{y}{x}$
- $g(x,y) = \frac{\sqrt{x}}{4y}$
- $g(x,y) = \sqrt{x} - y^3$

الدوال الهامشية للمتغيرات العشوائية المشتركة:

- من خلال دوال الاحتمال المشتركة للمتغيرين العشوائيين x, y يمكننا إيجاد الدوال الهامشية الخاصة بكل من x, y.

الجداول الهامشية في حالة المتغيرات العشوائية المتقطعة الثنائية:

$$p(X = x) = \sum_{y} p(X = x, Y = y)$$

$$p(Y = y) = \sum_{x} p(X = x, Y = y)$$

الدوال الهامشية في حالة المتغيرات العشوائية المتصلة الثنائية:

$$f(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

مثال(*):

ليكن:

$$f(x,y)=\begin{cases}\frac{1}{3}(x+y); 0<x<1, 0<y<2\\ 0; O.W.\end{cases}$$

فاحسب $f(x), f(y)$

مثال(**):

ليكن لدينا جدول التوزيع الاحتمالي التالي:

		X		
		0	**1**	**2**
Y	**0**	0	1/18	2/18
	1	1/18	2/18	3/18
	2	2/18	3/18	4/18

فاحسب $p(X = x), p(Y = y)$

التباين للمتغيرات العشوائية المشتركة:

تأخذ الصيغة $cov(x,y)$ حيث x ,y متغيرين عشوائيين و يسمى التباين المشترك وتأخذ الصيغة التالية:

$$cov(x,y) = E(xy) - E(x)E(y)$$

- حيث يتم حساب $E(xy)$ من خلال التوقع الرياضي المشترك باعتبار $g(x,y) = xy$
- ويتم حساب $E(x)$ و $E(y)$ من خلال دوالهما الهامشية.
- وبالتالي فان التباين والتباين المشترك للمتغيرين العشوائيين x, y تصبح مصفوفة ثنائية قطرها يمثل التباين لكل منهما والقيم اعلى واسفل القطر تمثل تباينهما المشترك.

$$V = \begin{bmatrix} var(x) & cov(x,y) \\ cov(y,x) & var(y) \end{bmatrix}$$

<u>مثال:</u>

أوجد التباين المشترك للمتغيرين العشوائيين المتصلين x, y في المثال(*) السابق؟

مثال:

أوجد التباين المشترك للمتغيرين العشوائيين المتقطعين x, y في المثال(**) السابق؟

استقلال المتغيرات العشوائية:

يقال عن المتغيرين العشوائيين **<u>المتقطعين</u>** X, Y بأنهما مستقلين اذا تحقق الشرط التالي:

$$p(X = x, Y = y) = p(X = x)\, p(Y = y)$$

لجميع قيم X و Y.

ويقال عن المتغيرين العشوائيين **<u>المتصلين</u>** X, Y بأنهما مستقلين اذا تحقق الشرط التالي:

$$f(x, y) = f(x)f(y)$$

لجميع قيم X و Y.

مثال :

لنفرض أن لدينا جدول التوزيع الاحتمالي المشترك التالي:

وضح اذا كان المتغيرين العشوائيين X و Y متغيرين عشوائيين مستقلين؟ ولماذا؟

X \ Y	1	2
0	1/10	2/10
1	0	1/10
2	3/10	3/10

مثال:

ليكن لدينا جدول التوزيع الاحتمالي التالي:

		X		
		0	1	2
Y	0	0	1/18	2/18
	1	1/18	2/18	3/18
	2	2/18	3/18	4/18

هل المتغيرين العشوائيين X و Y متغيرين عشوائيين مستقلين؟ ولماذا؟

<u>مثال :</u>

ليكن:

$$f(x,y)=\begin{cases}4xy; 0<x<1, 0<y<1\\ 0; O.W.\end{cases}$$

هل المتغيرين العشوائيين X و Y متغيرين عشوائيين مستقلين؟ ولماذا؟

مثال:

ليكن:

$$f(x,y)=\begin{cases}\frac{1}{3}(x+y);0<x<1,0<y<2\\ 0;O.W.\end{cases}$$

هل المتغيرين العشوائيين X وY متغيرين عشوائيين مستقلين؟ ولماذا؟

<u>مثال :</u>

ليكن:

$$f(x,y)=\begin{cases}2e^{x+y}; 1<x<2, 1<y<4\\ 0; O.W.\end{cases}$$

ادرس استقلالية المتغيرين العشوائيين X وY؟

بعض النتائج المهمة:

- من أجل أي متغيرين عشوائيين X و Y فان:

$$E(x+y) = E(x) + E(y)$$

وبشكل عام:

$$E(xy) \neq E(x)E(y)$$

الا في حالة خاصة وذلك عندما يكون المتغيرين العشوائيين X و Y متغيرين عشوائيين مستقلين فإن:

$$E(xy) = E(x)E(y)$$

- في حالة كون X و Y مستقلين فإن:

$$cov(x,y) = 0$$

أثبت؟

التوزيع الاحتمالي الشرطي:

- اذا كان X, Y متغيرين عشوائيين متقطعين، عندئذ نسمي $p(x|y)$ احتمال وقوع x مع العلم ان y قد وقع مسبقاً. وقانونه:

$$p(x|y) = \frac{p(X = x, Y = y)}{p(Y = y)}$$

$$p(y|x) = \frac{p(X = x, Y = y)}{p(X = x)}$$

- اذا كان X, Y متغيرين عشوائيين متصلين، عندئذ نسمي $f(x|y)$ احتمال وقوع x مع العلم ان y قد وقع مسبقاً. وقانونه:

$$f(x|y) = \frac{f(x, y)}{f(y)}$$

$$f(y|x) = \frac{f(x, y)}{f(x)}$$

مثال :

لنفرض أن لدينا جدول التوزيع الاحتمالي المشترك التالي:

كون جدول التوزيع الاحتمالي الشرطي لـ:

- $p(x|y=3)$
- $p(y|x=1)$
- $p(X>1|y=3)$
- $p(X\leq 1|y=3)$

		X		
		0	**1**	**2**
Y	**2**	3/9	1/9	2/9
	3	1/9	0	2/9

مثال :

ليكن:

$$f(x,y)=\begin{cases}\frac{1}{3}(x+y);0<x<1,0<y<2\\ 0;O.W.\end{cases}$$

أوجد $f(x|y)$ وكذلك $f(y|x)$.

نتيجة:

اذا كان X و Y متغيرين عشوائيين متصلين ومستقلين فان:

$$f(x|y) = f(x)$$

$$f(y|x) = f(y)$$

البرهان:

تمرين:

ليكن:

$$f(x,y)=\begin{cases}\frac{1}{6};0<x<2,0<y<3\\ 0;O.W.\end{cases}$$

أوجد $f(x|y)$ وكذلك $f(y|x)$.

معامل الارتباط بين المتغيرات العشوائية:

- يرمز له بالرمز $R_{x,y}$ ويعرف بالقانون التالي:

$$R_{x,y} = \frac{cov(x,y)}{\sigma_x \ \sigma_y}$$

حيث σ_x,σ_y يمثلان الانحراف المعياري للمتغيرين العشوائيين X وY.

- قيمة معامل الارتباط $R_{x,y}$ تكون في الفترة:

$$-1 \leq R_{x,y} \leq 1$$

- كلما اقتربت قيمة $R_{x,y}$ من ± 1 كان الارتباط قوي.
- عندما تكون قيمة $R_{x,y}$ موجبة كان الارتباط طردي.
- عندما تكون قيمة $R_{x,y}$ سالبة كان الارتباط عكسي.
- كلما اقتربت قيمة $R_{x,y}$ من الصفر كان الارتباط ضعيف.
- عندما: $R_{x,y} = 0 \quad \leftarrow$ x,y مستقلين.
- عندما: $R_{x,y} = \pm 1 \quad \leftarrow$ ارتباط تام.

مثال :

لنفرض أن لدينا جدول التوزيع الاحتمالي المشترك التالي:

احسب معامل الارتباط $R_{x,y}$؟

	X = 0	X = 1
Y = 1	3/10	1/10
Y = 2	2/10	0
Y = 3	0	4/10

مثال :

ليكن:

$$f(x,y)=\begin{cases}4xy; 0<x<1, 0<y<1\\ 0; O.W.\end{cases}$$

احسب معامل الارتباط $R_{x,y}$؟

نظرية النهاية المركزية

نظرية النهاية المركزية:

لأى عينة كبيرة ($n \geq 30$) فإن توزيع المعاينة الإحصائية يئول إلى التوزيع الطبيعى بغض النظرعن التوزيع الإحتمالى للمجتمع.

- أى أنه إذا كان لدينا مجموعة من المتغيرات المستقلة $X_1, X_2, X_3, \ldots, X_n$ تمثل عينات عشوائية وكان

$$E(X_i) = \mu_i, \qquad var(X_i) = \sigma_i^2$$

فإذا عرفنا المتغير $X = X_1 + X_2 + X_3 + \ldots, + X_n$ فإن المتغير العشوائي Z يكون:

$$Z = \frac{X - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i}}$$

نتائج على نظرية النهاية المركزية:

- إذا كان لدينا مجموعة من المتغيرات المستقلة $X_1, X_2, X_3, \dots, X_n$ والتي تمثل عينات عشوائية وكان

$$E(X_i) = \mu, var(X_i) = \sigma^2; i = 1,2, \dots, n$$

فإذا عرفنا المتغير $X = X_1 + X_2 + X_3 + \dots, + X_n$ فإن التوقع الرياضى لهذا المتغير يساوي:

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = \mu + \mu + \cdots + \mu = n\mu$$

وكذلك التباين له يساوى:

$$var(X) = var(X_1) + var(X_2) + \cdots + var(X_n) = \sigma^2 + \sigma^2 + \cdots + n\sigma^2$$

وبتطبيق نظرية النهاية المركزية على المتغير X فإن:

$$Z = \frac{X - n\mu}{\sigma\sqrt{n}}$$

نتائج على نظرية النهاية المركزية:

- إذا عرفنا المتغير $\bar{X}$ بأنه:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

فان توقعه يكون:

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}\left(E(X_1) + E(X_2) + \cdots + E(X_n)\right) = \frac{1}{n}(\mu + \mu + \cdots + \mu)$$

$$= \frac{n\mu}{n} = \mu$$

وتباينه يكون:

$$var(\bar{X}) = \frac{\sigma^2}{n}$$

أثبت؟

نتائج على نظرية النهاية المركزية:

- إذا كان لدينا مجتمع له الوسط الحسابى μ وانحرافه المعيارى σ ، أخذت منه عينة حجمها n أي $X_1, X_2, X_3, \dots, X_n$ لها الإنحراف المعيارى S فإن توزيع العينه الإحصائي $\bar{X}$ له الخصائص الآتية:
- إذا كان المجتمع له توزيع طبيعي ومتوسطه μ **وإنحرافه المعيارى σ معلوم** فإن:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), Z = \frac{\bar{X} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}, \qquad Z \sim N(0,1)$$

- إذا كان المجتمع له توزيع طبيعي أو غير ذلك **وإنحرافه المعيارى σ مجهول وحجم العينة كبير أي أن n > 30** فإن:

$$\bar{X} \sim N(\mu, \frac{S^2}{n}), \quad Z = \frac{\bar{X} - \mu}{\left(\frac{S}{\sqrt{n}}\right)}, \quad Z \sim N(0,1)$$

- إذا كان المجتمع له توزيع طبيعي أو غير ذلك **وإنحرافه المعيارى σ مجهول وحجم العينة صغير أي أن n < 30** فإن:

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$$

حيث t(n-1) هو توزيع الطالب t-distribution بدرجة حرية v=n-1

أمثلة:

- مثال (1):

مجتمع له توزيع طبيعى وسطه الحسابى 25 أختيرت منه عينة عشوائية حجمها 50 وإنحرافها المعيارى 5 أوجد:

$$p(25.1 < \bar{X} < 25.9)$$

الحل:

$$p(25.1 < \bar{X} < 25.9) = p\left(\frac{25.1-25}{\frac{5}{\sqrt{50}}} < \frac{\bar{X}-25}{\frac{5}{\sqrt{50}}} < \frac{25.1-25}{\frac{5}{\sqrt{50}}}\right) = p(0.14 < Z < 1.27) = p(Z < 1.27) - p(Z < 0.14) = 0.8980 - 0.5557$$

$$= 0.3423$$

عن طريق برنامج R:

```
> pnorm(25.9, mean = 25, sd = 0.707) - pnorm(25.1, mean = 25, sd = 0.707)
[1] 0.3422483
```

أمثلة:

- مثال (2):

أخذت عينة حجمها 25 من مجتمع طبيعي له الوسط 70 وانحراف معيارى 10 أوجد:

$$p(66 < \bar{X} < 74)$$

عن طريق برنامج R:

```
> pnorm(74, mean = 70, sd = 2) - pnorm(66, mean = 70, sd = 2)
[1] 0.9544997
```

تقريب توزيع ذي الحدين بالتوزيع الطبيعي:

اذا كان X يتبع توزيع ذي الحدين (ثنائي الحد) فان قانونه:

$$f(x) = p(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- في حالة n صغيرة نستخدم توزيع ذي الحدين.
- في حالة n كبيرة نستخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين حيث:

$$\mu = \ n\,p, \qquad \sigma^2 = n\,p\,(1-p)$$

$$X \sim N\left(\mu = n\,p, \sigma^2 = n\,p\,(1-p)\right)$$

تقريب توزيع ذي الحدين بالتوزيع الطبيعي:

مثال 1:

لنأخذ التوزيع الثنائى فى حالة n = 80 و p=0.35

احسب $p(X \geq 30)$ باستخدام التوزيع الثنائى ثم باستخدام التوزيع الطبيعي؟

الحل عن طريق برنامج R:

```
#Using binomial distribution:
> pbinom(80, size = 80, prob = 0.35) - pbinom(29, size = 80, prob = 0.35)
[1] 0.3588295
#Approximating the Binomial distribution
#Using the normal distribution:
> 1-pnorm(30, mean = 28, sd = 4.26)
[1] 0.319362
#Using continuity correction:
> 1-pnorm(29.5, mean = 28, sd = 4.26)
[1] 0.3623769
```