الفصل الرابع: التوزيعات الاحتمالية

1. التوزيعات الاحتمالية المتقطعة

بعد تبيان مفهوم المتغيرات العشوائية والتوزيع الاحتمالي يمكننا الآن أن نتطرق لعدد من التوزيعات الاحتمالية الشهيرة، تستخدم هذه التوزيعات في حل العديد من المسائل في مجال التسيير الصناعي والتجاري وفي الإدارة. ومن أكثر هذه التوزيعات شيوعا: التوزيع ذي الحدين (الثنائي) وتوزيع بواسون. في النهاية يفترض أن يكون الطالب قادرا على استذكار القوانين المدروسة وخصائصها الأساسية، ومعرفة متى وكيف يمكن استخدام كل قانون.

1.1. التوزيع المنتظم المتقطع

هو توزيع الاحتمالات حيث يكون عددًا محدودًا من القيم المتباعدة بالتساوي ويمكن ملاحظتها بشكل متساوٍ تقريبًا؛ فكل قيمة من القيم n يكون لها احتمال متساوٍ مع $1/n$. وبمعنى آخر فإن "التوزيع المنتظم المتقطع" سيكون "عددًا معرفًا من النتائج المتباعدة بالتساوي والتي لها نفس نسبة احتمال الحدوث."

ليكن n عدد طبيعي موجب، نقول عن المتغير العشوائي X أنه يتبع التوزيع المنتظم في المجال $[1-n]$ اذا كانت قيمه الممكنة 1، 2، ... n مع الاحتمالات $p(X=k)=\frac{1}{n}$

وعليه يكون

$$\sum_{k=1}^{n}\frac{1}{n}=1$$

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

$$V(X)=\frac{n^2-1}{12}$$

هذا القانون يسمح بنمذجة الظواهر ذات نتيجة عشوائية تأخذ قيم طبيعية محصورة بين $[1-n]$.

2.1. توزيع برنولي

نقول عن تجربة أنها تتبع توزيع برنولي[1] إذا كانت تحتمل نتيجتين (حدثين) متنافيتين A و'A. نسمي A نجاح و'A فشل.

نعتبر المتغير عشوائي X التي تمثل **عدد مرات النجاح**، تأخذ X القيمة 1 عند تحقق الحدث A و0 في الحالة المعاكسة.

نرمز عادة ب p "احتمال النجاح" لاحتمال تحقق الحدث A

---

[1] بإسم Jacques Bernoulli الذي درس هذا التوزيع في أواخر القرن 17.

و$q = 1 - p$ احتمال الحدث المعاكس (الفشل). يعين توزيع برنولي كما يلي :

$$\boxed{P(X=1)=p, \qquad P(X=0)=q, \qquad X=0,\,1.}$$

ونكتب **$X \sim B(1, p)$**

خصائص توزيع برنولي

$E(X) = \Sigma x_i p_i = 1.p + 0.q = p \Rightarrow \quad E(X) = p.$

$V(X) = E(X^2) – E(X)^2 = (1^2.p + 0^2.q) – p^2 = p – p^2 = p(1- p) = pq \Rightarrow \quad$ **$V(X) = qp.$**

3.1. توزيع ثنائي الحد

يستخدم هذا التوزيع لإيجاد الاحتمالات في الحالات التي يكون للتجربة العشوائية ذات نتيجتان فقط (النجاح أو الفشل). حيث إذا كررنا تجربة برنولي n مرة فإن X (عدد مرات النجاح) تأخذ القيم:

$$X = 0, 1, 2, 3, \ldots n$$

لنفترض في التجربة لرمي قطعة نقدية مكررة عدد n من المرات، وX عدد مرات الحصول على صورة (F):

حالة : $n = 2$ $\quad X = 0, 1, 2.$

$P(X = 0) = q*q = q^2, \quad P(X=1) = P(FP) + P(PF) = p*q + q*p = \mathbf{2p^1q^1}$

حالة : $n = 3$ $\quad X = 0, 1, 2, 3.$

$P(X=3) = P(FFF) = p*p*p = p^3, P(X=2) = P(FFP \text{ ou } PFF \text{ ou } FPF) = \mathbf{3p^2q^1}$

حالة : $n = 4$ $\quad X = 0, 1, 2, 3, 4.$

$P(X=3) = P(FFFP \text{ ou } PFFF \text{ ou } FPFF \text{ ou } FFPF) = \mathbf{4\,p^3q^1}$

في النتيجة الأخيرة نلاحظ العدد 3 هو x ، العدد 1 هو $n-x$ ، والعدد 4 هو عدد الطرق الملائمة للحصول على ثلاث نجاحات من بين $(n=4)$ تجارب، ويمكن حسابه كما يلي:

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

وبالتالي فاحتمال عدد ما x من النجاحات من بين n تجربة برنولية يحسب كما يلي:

$$\boxed{P(X=x) = C_n^x p^x q^{n-x}, \qquad x = 0,1,2,3,\ldots,n \qquad , \qquad n = 1,2,3,\ldots}$$

حيث x عدد مرات النجاح، p احتمال النجاح في التجربة الواحدة (يبقى ثابت عند تكرار التجربة)، $q = 1-p$ احتمال الفشل وn عدد التجارب. و هو تعريف " قانون التوزيع الثنائي" ويكتب قانون التوزيع الاحتمالي أيضا كما يلي:

$$P(X=x) = C_n^x p^x (1-p)^{n-x}$$

ونكتب **$X \sim B(n, p)$**.

شروط استخدام التوزيع الثنائي

- تجربة برنولية مكررة عدد محدد من المرات
- احتمال النجاح في التجربة ثابت (التجارب مستقلة)

مثال 1: أحسب عند رمي قطعة نقدية متوازنة 4 مرات، ما هو احتمال الحصول على:

ولا مرة صورة، مرة واحدة، مرتين، 3 مرات، 4 مرات.

$P(X = x) = C^x_n \; p^x \, q^{n-x} \Rightarrow P(X = 0) = C^0_4 \; 0.5^0 \, 0.5^4 = 1/16$

$P(X = 1) = C^1_4 \; 0.5^1 \, 0.5^3$

$P(X = 2) = C^2_4 \; 0.5^2 \, 0.5^2$

مثال 2: نسحب بالإرجاع 3 كريات من صندوق يحتوي 5 كريات منها 3 حمراء.

أحسب احتمال الحصول على كريتين حمراء.

$P(X=2) = C^2_3 \; (3/5)^2 \, (2/5)^1$

خصائص التوزيع الثنائي

التوقع والتباين:

يمكن اعتبار X مجموع متغيرات مستقلة برنولية $X = X_1 + X_2 + \dots X_i + \dots + X_n$ لها نفس المعلمة p وبالتالي نفس التوقع $(E(Xi) = p)$ أيضا.

إذا باستخدام خصائص التوقع والتباين نجد:

$E(X) = E(X_1 + X_2 + \dots X_i + \dots + X_n) = \Sigma E(Xi) = \Sigma pi = n\,p \Rightarrow \mathbf{E(X) = np}$

$V(X) = V(X_1 + X_2 + \dots X_i + \dots + X_n),$

Xi مستقلة إذن:

$$V(X) = \Sigma V(Xi) = \Sigma pq \Rightarrow \mathbf{V(X) = npq}$$

مثال 3: إذا كانت نسبة نجاح الطلبة الذين يحضرون المحاضرة في مقياس الإحصاء هي 0.6 ، وإذا كان عدد الطلبة الذين يحضرون المحاضرة بانتظام هو 10 طلبة. فإذا افترضنا المتغير العشوائي يمثل عدد الطلبة الناجحين الذين يحضرون المحاضرة.

1 –حدد شكل دالة الاحتمال لهذا المتغير

2–احسب الاحتمالات التالية:

أ –ما هو احتمال نجاح **3** طلبة الذين يحضرون المحاضرة ؟

ب –ما هو احتمال نجاح طالب واحد على الأقل؟

ج –ما هو احتمال نجاح طالبين على الأكثر من الذين يحضرون المحاضرة ؟

3–احسب المتوسط الحسابي، والانحراف المعياري لعدد الطلبة الناجحين.

الحل:

1 –شكل دالة الاحتمال لهذا المتغير

$n = 10 \quad q = p - 1 = 0.4 \quad P = 0.6$ وبالتالي:

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

$$f(x) = c_{10}^X 0.6^x (0.4)^{10-x}$$

$$x = 0,1,2 \ldots \ldots \ldots \ldots \ldots .10$$

2–حساب الاحتمالات:

أ –حساب احتمال نجاح 3 طلبة الذين يحضرون المحاضرة:

$$P(x=3) = f(3)$$

$$f(3) = c_{10}^3 0.6^3 (0.4)^{10-3}$$

$$= 120 \times 0.216 \times 1.6384 \times 10^{-3} = 0.04$$

ب –حساب احتمال استجابة مريض واحد على الأقل:

$$P(x \geq 1) = 1 - f(0)$$

$$= 1 - c_{10}^0 0.6^0 (0.4)^{10-0}$$

$$= 1 - 1 \times 1 \times 1.048576 \times 10^{-4} = 0.99$$

ج –حساب احتمال نجاح طالبين على الأكثر من الذين يحضرون المحاضرة:

$$P(x \leq 2) = f(2) + f(1) + f(0)$$

$$= c_{10}^2 0.6^2 (0.4)^{10-2} + c_{10}^1 0.6^1 (0.4)^{10-1} + c_{10}^0 0.6^0 (0.4)^{10-0}$$

$$= 0.01$$

3–حساب التوقع الرياضي، والانحراف المعياري لعدد الطلبة الناجحين:

–التوقع الرياضي

$$E(x) = np = 10 \times 0.6 = 6$$

– التباين

$$V(x) = np(1-p) = 10 \times 0.6 \times 0.4 = 2.4$$

–الانحراف المعياري

$$\sigma(x) = \sqrt{np(1-p)} = \sqrt{2.4} = 1.54$$

4.1. توزيع بواسون

يستخدم هذا التوزيع لتحديد احتمال عدد معين من الوقائع او الأحداث او النجاحات في مدة زمنية أو في منطقة مكانية محددة، وذلك عندما تكون الأحداث مستقلة عن بعضها البعض ويبقى متوسطها (معدل وقوعها) ثابتا لوحدة زمنية او مكانية معينة، مثل عدد حوادث السيارات في منطقة مكانية، عدد الاخطاء الكتابية في صفحة، عدد المكالمات خلال ساعة او اسبوع او شهر ... الخ.

التوزيع الاحتمالي لمتغير بواسون والذي يعبر عن عدد مرات النجاح التي تحدث في مدة زمنية أو في منطقة مكانية محددة هو:

$$P(X=k)=\frac{\lambda^k \ e^{-\lambda}}{k!} \qquad ; x=0,1,2,\dots n$$

ونكتب: $X \sim P(\lambda)$

حيث أن:

k :عدد النجاحات .

$P(X=k)$: _ .احتمال عدد k من النجاحات

λ : متوسط عدد النجاحات في وحدة زمانية او مكانية.

e : .أساس نظام اللوغاريتم الطبيعي

يمتاز هذا التوزيع بتساوي الامل الرياضياتي مع التباين

$$\mathrm{E(X)=V(X)=\lambda}$$

مثال 4 : إذا كان عدد حوادث السيارات في مدينة معينة يتبع التوزيع البواسوني بمتوسط 5 حوادث خلال الأسبوع .إذا افترضنا أن X متغير عشوائي يمثل عدد حوادث السيارات خلال أسبوع .

1–حدد شكل دالة الاحتمال f(x) لهذا المتغير

2–أحسب الاحتمالات التالية:

أ –ما هو احتمال حدوث حادثين خلال أسبوع؟

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

ج –ما هو احتمال حدوث ثلاثة حوادث على الأكثر خلال أسبوع؟

3–احسب التوقع الرياضي، التباين والانحراف المعياري لعدد الحوادث.

4–حدد شكل التوزيع.

الحل:

1–شكل دالة الاحتمال:

بما أن متوسط عدد حوادث السيارات خلال أسبوع هو 5 وبالتالي تكون دالة الاحتمال هي توزيع بواسون كما يلي:

$$P(X=k)=\frac{5^k\ e^{-5}}{k!}\qquad ; x=0,1,2,\dots n$$

ونكتب: $X\sim P(5)$

2–حساب الاحتمالات:

أ –احتمال حدوث حادثين خلال أسبوع هو (من اجل2= k)

$$f(2)=\frac{5^2e^{-5}}{2!}=\frac{0.16}{2\times 1}=0.08$$

ب –احتمال حدوث حادث واحد على الأقل خلال أسبوع هو

$$f(X\geq 1)=1-f(X<1)=1-f(X=0)$$

$$=1-\frac{5^0e^{-5}}{0!}=0.99$$

ج –احتمال حدوث ثلاثة حوادث على الأكثر خلال أسبوع هو

$$f(X\leq 3)=f(3)+f(2)+f(1)+f(0)$$

$$=\frac{5^3e^{-5}}{3!}+\frac{5^2e^{-5}}{2!}+\frac{5^1e^{-5}}{1!}+\frac{5^0e^{-5}}{0!}=0.25$$

3–حساب التوقع الرياضي، والانحراف المعياري لعدد الحوادث:

- التوقع الرياضي

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

- التباين

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

- الانحراف المعياري

$$\sigma(x)=\sqrt{\lambda}=\sqrt{5}=2.23$$

4–تحديد شكل التوزيع:

التوزيع البواسوني دائما موجب الالتواء.

خلاصة : الجدول الملحق يلخص أهم النقاط حول التوزيعات المتقطعة الشهيرة.

التوزيع	متى يستخدم	القيم الممكنة للمتغيرة	الاحتمال	التوقع والتباين
برنولي **X~B(1, p)**	تجربة واحدة (غير مكررة) تقبل نتيجتين.	$X = \{0, 1\}$	$P(X = 1) = p$, $P(X = 0) = 1 - p = q$	$\mu = p, \quad \sigma^2 = pq$
الثنائي **X~B(n, p)**	تجارب ثنائية النتيجة، مكررة ومستقلة (p ثابت).	$X = \{0,1,2\ldots,n\}$	$P(X = x) = C^x_n\, p^x\, q^{n-x}$ $C_n^x = \dfrac{n!}{x!(n-x)!}$	$\mu = np, \sigma^2 = npq$
باسكال (الثنائي السالب)	X هي عدد التجارب اللازمة للحصول على عدد r من النجاحات في تجارب برنولية مكررة.	$X = \{r, r+1, r+2, \ldots, +\infty\}$	$P(X = x) = C^{r-1}_{x-1}\, p^r\, q^{x-r}$	$\mu = r/p$, $\sigma^2 = rq/p^2$
الهندسي	X هي عدد التجارب اللازمة للحصول على النجاح الأول في تجارب برنولية مكررة.	$X = \{1,2,\ldots,+\infty\}$	$P(X = x) = q^{x-1}p$	$\mu = 1/p$, $\sigma^2 = q/p^2$
التوزيع المتعدد	هو تعميم للتوزيع الثنائي على تجربة مكررة متعددة النتائج.	$\forall i, 0 \le xi \le Ni,$ $\sum_{i=1}^{k} xi = n,$ $\sum_{i=1}^{k} Ni = N$	$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \dfrac{n!}{x_1!x_2!\ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$	$E(X_k) = np_k$ $V(X_k) = np_kq_k$
بواسون **X~P(λ)** **λ > 0**	X عدد النجاحات في عدد كبير من التجارب البرنولية (عدد الوحدات التالفة في شحنة). أو عدد من الأحداث في فترة زمن.	$X = \{0,1,2,\ldots,+\infty\}$	$P(X = x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ $P(X = 0) = e^{-\lambda}$ $P(X \ge 1) = 1 - e^{-\lambda}$	$E(x) = V(x) = \lambda$

2. التوزيعات الاحتمالية المتصلة

1.2. التوزيع الأسي

عادة ما يستخدم التوزيع الأسي في مسائل متعلقة بقياس الزمن. من ذلك مدة خدمة شباك البريد، مدة مكالمة هاتفية، مدة تفريغ باخرة شحن، مدة تصليح آلة، مدة انتظار زبون قبل الحصول على الخدمة...في العلوم الدقيقة يستخدم التوزيع الأسي لتمثيل مدة حياة الذرات المشعة (atomes radioactives) قبل أن تتفكك، حيث يعبر الوسيط عن اللحظة التي يبقى فيها نصف المجتمع الأصلي.

من الضروري فهم الآتي: كقاعدة عامة يستخدم التوزيع الأسي لتمثيل مدة حياة ظاهرة ما إذا كان لها متوسط ثابت $1/\lambda$ وكانت هذه الظاهرة لا تخضع للتقادم (vieillissement) أي أن مدة حياة الظاهرة بعد لحظة ما T لا تتبع اللحظة T؛ 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 ليس مستقلا عن المدة التي عملتها الآلة من قبل، كذلك الأمر بالنسبة لمدة حياة الإنسان.

عمليا، نتحقق من دقة تمثيل التوزيع الأسي _أو أي توزيع آخر_ لظاهرة ما من خلال تقنيات اختبارات الفروض، وبالتحديد اختبار التجانس و التعديل.

نشير أخيرا إلى أن للتوزيع الأسي علاقة بالتوزيع 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 التوزيع بواسون فإن المدة الزمنية بين وصول زبون "أ" والزبون الموالي تتبع التوزيع الأسي. تتبين هذه العلاقة عند استنتاج صيغة القانون الأسي.

أ. دالة الكثافة والدالة التجميعية للتوزيع.

بينت دراسة أن عدد حوادث العمل في معمل معين تتبع توزيع بواسون بمعدل λ حادث يوميا.

أوجد احتمال أن يسجل حادث على الأقل (حادث أو أكثر) في مدة t يوم.

$P(X \geq 1) = 1- P(0) = 1-[\lambda^{0t} * e^{-\lambda t}/0!] \quad \Rightarrow \mathbf{P(X \geq 1) = 1-e^{-\lambda t}}$

لنرمز ب T للزمن (باليوم) بين حادثين إذن سيكون لدينا $f(t)$ دالة الكثافة للزمن بين حادثين، و $F(t) = P(T \leq t)$ دالة التوزيع ل T.

لنحسب احتمال P أن يكون الزمن بين حادثين يوم أو أقل:

لدينا $P = P(T \leq t = 1)$ إذن:

$P = F(t = 1)$ .... (1)

نلاحظ من ناحية أخرى أن P هو معادل لاحتمال أن يسجل على الأقل حادث في يوم معين:

$P = P(X \geq 1) = 1-e^{-\lambda t}$ ....(2)

من (1) و(2) نستنتج أن $\boxed{\mathbf{F(t) = 1 - e^{-\lambda t}}}$............(3)

و منه $f(t) = F(t)' = (1 - e^{-\lambda t})'$

إذن $\mathbf{f(t) = \lambda e^{-\lambda t}}$

قاعدة: إذا كان حدث عشوائي ما يتكرر في الزمن وفق توزيع بواسون:

$$\boxed{p_{\tau}(x) = \frac{(\lambda\tau)^{x} e^{-\lambda\tau}}{x!}}$$

فإن الزمن t بين حادثين يتبع التوزيع التالي:

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & ; t > 0 \\ 0 & ; t \le 0 \end{cases}$$

حيث λ عدد حقيقي موجب. ويسمى هذا التوزيع التوزيع الأسي ويسمى أيضا التوزيع الأسي السالب لعلاقته بتوزيع بواسون.

ب. خصائص التوزيع الأسي

$$E(X) = \frac{1}{\lambda} \quad ; \quad V(X) = \frac{1}{\lambda^2}$$

كما يمتاز التوزيع الاسي بخاصية تسمى غياب الذاكرة، ليكن متغير عشوا X ئ ي يتبع التوزيع الاسي

$$X \sim \xi(\lambda)$$

اذن:

$$p(X > t + s / X > s) = \frac{p(x > t + s /\ \ X > s)}{p(X > s)} = \frac{p(X > t + s)}{p(X > s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}}$$
$$= e^{-\lambda(t+s-s)} = e^{-\lambda t} = p(X > t)$$

2.2. التوزيع الطبيعي

يعد التوزيع الطبيعي[2] أو توزيع لابلاس قوس 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 شارع 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.

أ. دالة الكثافة

لو 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 (الشكل 1 و3).

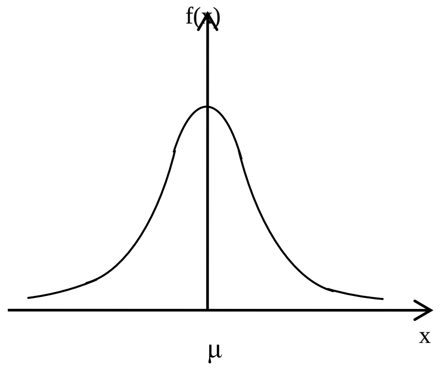
تكتب دالة الكثافة لمنحنى للتوزيع الطبيعي كما يلي

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}exp\left\{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}\qquad ;\forall x\in R$$

حيث μ و σ هما على التوالي التوقع والانحراف المعياري.

ونكتب: $X\sim\mathbb{N}(\mu,\sigma)$

تمثيلها البياني التقريبي في الشكل 1 المقابل



الشكل: 1 الشكل العام للتوزيع الطبيعي

دالة التوزيع (الدالة التجميعية) للتوزيع الطبيعي تكتب كما يلي:

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

استخدام المتغير المعياري: يستخدم المتغير المعياري Z = (X-μ)/σ لتكوين الجداول الإحصائية للاحتمالات:

[2] باسم العالمان الرياضيان الفيزيائيان و الفلكيان الفرنسي Pière Simon de Laplace (1749-1827) والألماني Carl Freidrich Gauss – الصورة لهذا الأخير-، (1777-1855) الذين كانا من أوائل من اكتشف هذا القانون. أما من أعطاه تسمية التوزيع الطبيعي فهو Pearson في 1893. أنظر J. J. Droesbeke (1997)، ص329 .

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

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

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

$$F(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2}} du$$

بالنظر إلى العلاقة الخطية بين المتغيرتين X و Z، فإن Z تتبع نفس توزيع X أي التوزيع الطبيعي. ونعلم أن:

$$E(z) = 0 \qquad ; \qquad V(z) = 1$$

ب. خصائص التوزيع الطبيعي

من خصائص التوزيع الطبيعي أنه يعتبر معتدلا لا مدببا ولا مفلطحا، حيث يعتبر معامل التفلطح $\alpha_4 = 3$ للتوزيع الطبيعي معيارا لاعتدال المنحنيات.

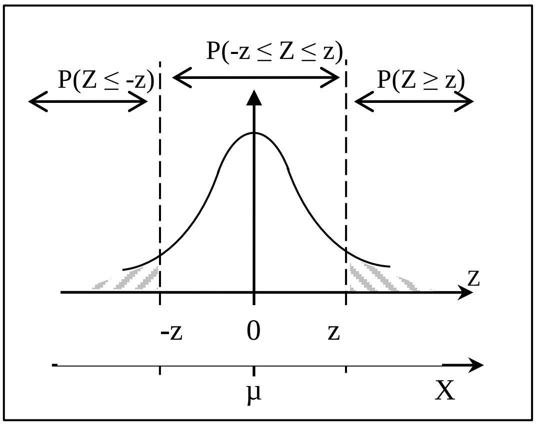
من خصائص التوزيع الطبيعي أيضا أنه متماثل حول $\alpha_3 = \frac{\mu_3}{\sigma^3} = 0$ القيمة المتوقعة

تماثل منحنى X حول المتوسط (أنظر الشكل 3) يعني تماثل لمنحنى Z حول 0، مما يعني أنه من أجل أي قيمة للمتغيرة المعيارية

: $z > 0$

$$P(0 \le Z \le z) = P(-z \le Z \le 0) = P(-z \le Z \le z) / 2$$

$$P(Z \le -z) = 1 - P(Z \le z) = P(Z \ge z)$$



الشكل 2: استخدام تماثل الوزيع الطبيعي في حساب الاحتمالات

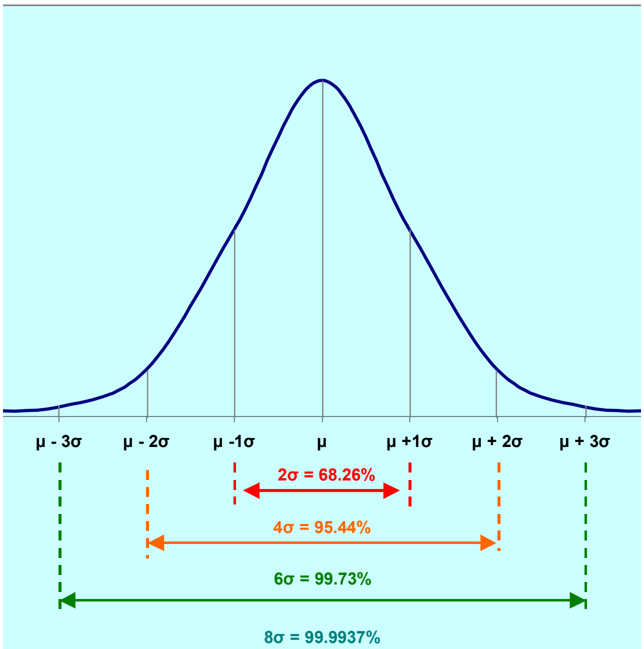
و لقد تم باستخدام المتغيرة المعيارية Z حساب الاحتمالات (المساحات) تحت المنحنى ومنها خاصة:

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

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

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

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



رسم 2 المساحات الأساسية تحت منحنى التوزيع الطبيعي

مثال1: باستعمال الجداول الاحصائية

1) أحسب : $P(0 \leq Z \leq z)$ حيث $z =1, 2, 3$

2) أحسب $P(-z \leq Z \leq z)$ من أجل نفس القيم ل z.

الحل:

1) 0.3413 ، 0.47725 ، 0.49865

2) 0.6827، 0.9545، 0.9973

مثال 2 . إذا كانت علامات الطلبة في قسم معين تتبع التوزيع الطبيعي بمتوسط حسابي يساوي10 وتباين يساوي4

1–حدد معالم التوزيع الاحتمالي لعلامات الطلبة؟

2–حدد شكل دالة كثافة الاحتمال؟

3–احسب احتمال أن تكون علامات الطلبة بين 8 و12 ؟

4–ما هو احتمال أن تكون علامات الطلبة أقل أو يساوي6 ؟

الحل:

1–تحديد معالم التوزيع الاحتمالي لعلامات الطلبة (نرمز لها بـ X)

بما أن علامات الطلبة هي متغير عشوائي يتبع التوزيع الطبيعي فإن معالمه تكون كالأتي:

التباين =4 الوسط الحسابي=10 وبالتالي:

$$\mu = 10 \quad ; \quad \sigma^2 = 4 \Longrightarrow \sigma = 2$$
$$X \sim \mathbb{N}(10; 2)$$

2–تحديد شكل دالة كثافة الاحتمال:

$$f(X) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{(x-10)^2}{4})}$$

3–احتمال أن تكون علامات الطلبة بين 8 و 12 هو

$$p(8 < X < 12) = p(8 - 10 < X - 10 < 12 - 10)$$
$$= p\left(\frac{-2}{2} < z < \frac{2}{2}\right) = 2p(0 < z < 1)$$

نستخدم جدول التوزيع الطبيعي المعياري لحساب الاحتمال

حيث نجد ان :

$$p(0 < z < 1) = 0.3413$$

وعليه

$$p(8 < X < 12) = 2(0.3413) = 0.6826$$

4–احتمال أن تكون علامات الطلبة أقل أو يساوي 6 هو

$$p(X \leq 6) = p(X - 10 \leq 6 - 10)$$
$$= p\left(z \leq \frac{-4}{2}\right) = p(z \leq -2) = p(z > 2) = 1 - p(z \leq 2)$$
$$= 1 - [0.5 + p(0 < z < 2)] = 0.5 - p(0 < z < 2)$$

نستخدم جدول التوزيع الطبيعي المعياري لحساب الاحتمال

حيث نجد ان : $p(0 < z < 2) = 0.4772$

$$p(X \leq 6) = 0.5 - 0.4772 = 0.0228$$

ج. العلاقة بين التوزيع الطبيعي والتوزيع الثنائي

في حالة n كبيرة و p غير قريب من 0 يمكن اعتبار التوزيع الثنائي كتقريب جيد للتوزيع الطبيعي. ويعطي التوزيعان نتائج أكثر تقاربا كلما كانت n كبيرة أكثر. ونكتب:

$$\lim_{n \longrightarrow \infty} P(a \leq z \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-z^2/2} dz, \qquad z = \frac{x - np}{\sqrt{npq}}$$

ويسرع تقارب التوزيع الثنائي من التوزيع الطبيعي كون p قريب من 0.5.

قاعدة التقريب:

- عموما نعتبر التقريب إلى التوزيع الثنائي ملائما عندما np و nq كلاهما أكبر من 5.
- عدد من الاحصائيين[3] يعتمد قاعدة أخرى هي أن يكون أحد الشرطين التاليين متوفرين:

$$npq \geq 9 \qquad ou \qquad n \geq 20, np \geq 10, nq \geq 10$$

[3] أنظر دراوزبيك 1997. ص 262.