كلية العلوم الاجتماعية- قسم الديموغرافيا

مادة الإحصاء الرياضي وتطبيق الاختبارات الإحصائية

ماستر 1 ديموغرافيا اجتماعية

د. راشدي خضرة

المحاضرة 6 و7 : المتغيرات العشوائية والتوزيعات الاحتمالية

الهدف من المحاضرة: يتعرف الطالب على أهم التوزيعات الاحتمالية وكيفية تطبيقها في الظواهر التي يدرسها والمرتبطة بتخصصه.

تمهيد:

1- المتغير العشوائي:

تعريف المتغير العشوائي:

يعرف المتغير العشوائي بأنه المتغير المجهول القيمة، وهو المتغير الذي يستخدم عند دراسة الظواهر العشوائية وكمثال بسيط على الظواهر العشوائية" رمي حجر النرد" حيث النتائج العددية لهذه الظواهر يعبر عنها بالمتغير العشوائي.

ويأخذ المتغير العشوائي غالبا قيما مختلفة ضمن مجال مستمر، حيث يعتبر هذا المتغير قابل للقياس، حيث يرمز لقيمه بأحرف معينة، إلا أنه غالبا ما يكون أرقام وقيم عددية مختلفة. وتعتبر المتغيرات العشوائية مختلفة تماما عن المتغيرات الجبرية، التي يمكن حسابها من المعادلات الرياضية بشكل مباشر. ويرمز للمتغير العشوائي بشكل عام بحرف من الحروف الأبجدية الكبيرة....,X, Y, Z ويرمز للقيم التي يأخذها المتغير بالحروف الأبجدية الصغيرة، $x, y, z, \dots$

أنواع المتغير العشوائي: يأخذ المتغير العشوائي نوعان:

1- المتغير العشوائي المنفصل(المتقطع):

يكون المتغير العشوائي منفصلا عندما يأخذ عددا كبيرا من القيم، قد تكون هذه القيم غير منتهية، ومن الممكن أن تكون منتهية، إلا أنها قابلة للعد. وتكون هذه القيم مختلفة عن بعضها البعض.

ويرمز لشكل مجموعة الاحتمالات تبعا للمتغير العشوائي المنفصل بالشكل التالي:

x = { X1, X2, X3,…Xi}

أمثلة عن المتغير العشوائي المنفصل:

– عدد الأولاد الذكور في الأسرة المكونة من أربع أولاد *X*، *X:{x=0,1,2,3,4}*

–عدد الوفيات في مصحة ما.

– عدد الزيجات المسجلة في مدينة ما

2– المتغير العشوائي المتصل (المستمر):

يكون المتغير العشوائي متغيرا متصلا عندما يأخذ عددا غير محدود أو لا نهائي من القيم داخل نطاق تغيره ويأخذ هذا النوع من المتغيرات الشكل التالي:

x ={ x:a => x=> b ; a, b € R}

وتحسب الاحتمالات للمتغير المتصل باستخدام المساحات تحت منحنى الدالة أو المتغير (مثل المنحى الطبيعي) كما سنرى لاحقا. أي تعبر على كثافة الاحتمال في تجربة معينة، حيث تأخذ احتمال للفترات أيضا، حيث يجب أن يكون لأي نقطة مهما كانت صغيرة متغير عشوائي مستمر.

أمثلة على المتغير العشوائي المتصل:

– أعمار الاشخاص.

– مدة الزواج

–الدخل الشهري

2- التوزيعات الاحتمالية:

2-1- تعريف التوزيع الاحتمالي

يرتبط مفهوم التوزيع الاحتمالي بمفهوم المتغير العشوائي، حيث يمكن للمتغيرات العشوائية أن تأخذ توزيعات احتمالية مختلفة، حيث يقوم التوزيع الاحتمالي بوصف الظواهر التي تحتوي متغيرات عشوائية.

ويأخذ التوزيع الاحتمالي شكل جدول أو دالة رياضية أو معادلة يبين كيف تتوزع الاحتمالات (والتي مجموعها الواحد الصحيح) على قيم المتغير المختلفة. أي أن التوزيع الاحتمالي والذي يسمى أيضاً دالة الاحتمال هو التوزيع أو الدالة التي تعطي احتمالات أن يأخذ المتغير العشوائي القيم المختلفة له.

وتنقسم التوزيعات الاحتمالية إلى نوعين: توزيعات منفصلة وأخرى متصلة وسنتناول في هذا المحور هذه التوزيعات مع الاخذ بعين الاعتبار أهمها والتي ترتبط بتخصص علم السكان.

2-2- التوزيعات الاحتمالية المنفصلة :

فإذا كان المتغير العشوائي محل الدراسة متغيراً متقطعاً مثل عدد الاطفال أو حجم الأسرة، كان التوزيع الاحتمالي له متقطعاً.

والتوزيعات الاحتمالية المتقطعة هي مجموعة القيم الممكنة مع الاحتمالات المرتبطة بقيم المتغيرة. نرمز للمتغيرة بحرف كبير وللقيم التي تأخذها المتغيرة بحرف صغير. نعبر عن احتمال قيمة معينة كما يلي $P(X = x)$ ونكتب أيضا : $f(x)$. وتسمى الدالة $f(x)$ دالة الكثافة الاحتمالية.

مثال: التوزيع الاحتمالي لم ع للمثال الأول (إلقاء مكعب نرد) يكتب كما يلي:

X	1	2	3	4	5	6	
P(X = x)	1/6	1/6	1/6	1/6	1/6	1/6	1

مثال 2. التوزيع الاحتمالي ل X، عدد مرات الصورة في رميتين لقطعة نقدية:

X	0	1	2	
P(X = x)	1/4	2/4	1/4	1

شروط دالة الكثافة للمتغيرة المتقطعة

نعبر عن احتمال قيمة معينة كما يلي $P(X = x)$ ونكتب أيضا : $f(x)$ وتسمى الدالة $f(x)$ دالة الكثافة الاحتمالية. لكي يمكن اعتبار دالة ما، أيا كانت، دالة كثافة احتمالية يجب أن يتحقق شرطان اثنان:

1) $f(x) \geq 0$

2) $\sum_x f(x) = 1$

مثال: نأخذ دالة الكثافة ل X نتيجة لإلقاء حجر نرد:

$$f(1) = f(2) = f(3) = \dots f(6) = 1/6 \geq 0$$

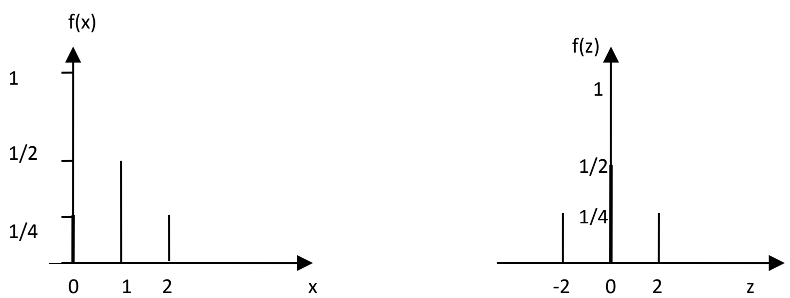
الشرط الأول محقق، والشرط الثاني أيضا لأن:

$$\Sigma f(x) = 1/6 + 1/6 + \dots + 1/6 = 6(1/6) = 1$$

التمثيل البياني لدالة الكثافة الاحتمالية للمتغير العشوائي المنفصل

تمثل المتغيرة العشوائية المتقطعة ليس من خلال منحنى ولكن من خلال أعمدة متوازية على محور X.

مثال: نمثل بيانيا منحنيات دوال الكثافة ل X و Z المعرفة على إلقاء قطعة نقدية مرتين.



التمثيل البياني لدالة الكثافة للمتغيرة العشوائية المتقطعة

دالة التوزيع F(x) للمتغير العشوائي المنفصل

تعرف دالة التوزيع – وتسمى أيضا "الدالة التجميعية" – كما يلي:

$$\mathbf{F(x) = P(X \leq x)}$$

ويمكن استنتاج دالة التوزيع من دالة الكثافة الاحتمالية f(x) كما يلي:

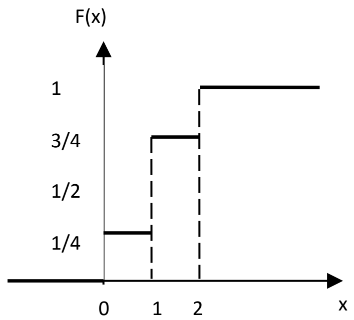
$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u)$$

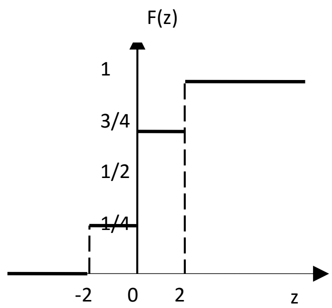
إذا كانت X تأخذ عددا منتهيا من القيم فإن F(x) يمكن تعريفها كما يلي:

$$F(x)=\begin{cases}0 & , \quad -\infty < x < x_1 \\ f(x_1) & , \quad x_1 \le x < x_2 \\ f(x_1)+f(x_2) & , \quad x_2 \le x < x_3 \\ \cdots & \\ f(x_1)+f(x_2)+\ldots+f(x_n), & \quad x_n \le x < +\infty \end{cases}$$

مثال:

أوجد قيم F(x) و F(z) للأمثلة السابقة ومثلهما بيانيا.





X	0	1	2
f(x)	1/4	1/2	1/4
F(x)=P(X≤x)	1/4	3/4	1

Z	−2	0	2
f(x)	1/4	1/2	1/4
F(x)=P(X≤x)	1/4	3/4	1

التمثيل البياني لدالة التوزيع للمتغيرة العشوائية المتقطعة

وهناك عدة توزيعات احتمالية متقطعة أهمها:

1-توزيع برنولي Distribution de Bernoulli

نقول عن تجربة أنها "برنولية" إذا كانت تحتمل نتيجتين (حدثين) متنافيتين A و A'. نسمي A نجاح و A' فشل.

نعتبر المتغيرة ع X التي تمثل **عدد مرات النجاح**، تأخذ X القيمة 1 عند تحقق الحدث A و 0 في الحالة المعاكسة.

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

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

$$V(X) = E(X^2) - E(X)^2 = (1^2.p + 0^2.q) - p^2 = p - p^2 = p(1-p)$$

$$= pq \Rightarrow \quad \mathbf{V(X) = qp.}$$

2-توزيع ذى الحدين

هو أحد التوزيعات الاحتمالية المنفصلة . ويستخدم لإيجاد احتمال وقوع حدث معين (نجاح) عدداً من المرات مقداره **X** من بين **n** من المحاولات لنفس التجربة (ونرمز لهذا الاحتمال بالرمز **(P(X)** عندما تتحقق الشروط التالية :

(1) هناك فقط ناتجان ممكنان ومتنافيان لكل محاولة

(2) المحاولات وعددها **n مستقلة** عن بعضها البعض

(3) احتمال وقوع الحدث المعين فى كل محاولة (النجاح) **P** ، ثابت ولا يتغير من محاولة لأخرى .

فيكون

$$P(X) = \frac{n!}{X!\,(n-X)!} P^x (1-P)^{n-X}$$

و بما أن :

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

يكون:

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

ويكون متوسط توزيع ذى الحدين(التوقع) هو:

$$\mu = np$$

وانحرافه المعيارى هو :

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

فإذا كانت **$p=1-p=0.5$** ، فإن توزيع ذى الحدين يكون متماثلا وإذا كانت **$p<0.5$** ، يكون التوزيع ملتوياً الى اليمين ، وإذا كانت **$p>0.5$** ،يكون التوزيع ملتوياً إلى اليسار.

مثال:

اذا كانت نسبة الامية في احدى المدن هي 0.32 ، تم اختبار 8 اشخاص من هذه المنطقة ، ماهو احتمال ان يكون :

1. 6 اشخاص اميون
2. كلهم اميون
3. و لا احد منهم امي
4. احدهم على الاقل امي
5. احدهم على الاكثر امي
6. 2 اميون على الاقل
7. 2 اميون على الاكثر

الحل:

لدينا:

$P = 0.32 \Rightarrow q = 0.68$, $n = 8$

1- 6 اشخاص اميون

$P(X = x) = C^x_n \ p^x q^{n-x} \Rightarrow P(X = 6) = C^6_8 \ 0.32^6 \ 0.68^2 = 0.000248$

2-كلهم اميون

$P(X = x) = C^x_n \ p^x q^{n-x} \Rightarrow P(X = 8) = C^8_8 \ 0.32^8 \ 0.68^0 = 0.000109$

3-ولا أحد منهم أمي

$P(X= x) = C^x_n \ p^x q^{n-x} \Rightarrow P(X = 0) = C^0_8 \ 0.32^0 \ 0.68^8 = 0.04571$

4-أحدهم على الأقل أمي:

$P(X \geq x) = P(X \geq 1) \Rightarrow P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8) = 1 - P(X = 0) = 1 - 0.04571 = 0.95429$

5- احدهم على الأكثر أمي

$$P(X \leq x) = P(X \leq 1) = P(X = 0) + P(X = 1)$$
$$= C^0_8\ 0.32^0\ 0.68^8 + C^1_8\ 0.32^1\ 0.68^7$$
$$= 0.04571 + 0.00119 = 0.045908$$

6- 2 اميون على الأقل

$P(X \geq x) = P(X \geq 2) \Rightarrow P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7) + P(X = 8) = 1 - P(X = 0) + P(X = 1) = 1 - 0.045908 = 0.954091$

7- 2 اميون على الأكثر

$$P(X \leq x) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= C^0_8\ 0.32^0\ 0.68^8 + C^1_8\ 0.32^1\ 0.68^7 + C^2_8\ 0.32^2\ 0.68^6$$
$$= 0.047296$$

مثال2:

- مثال 3 : احتمال شفاء مرضى من مرض نادر في الدم هو 0.4 ، اذا شخص 10 أفراد ووجد أنهم يحملون المرض ما هو احتمال :

1) أن يعيش 3
2) أن يعيش على الأقل 8
3) أن يعيش من 2 الى 5

الحل: لنفترض أن X هو عدد الأفراد الذين سيعيشون ، وبالاعتبار لجدول التوزيع ذو الحدين

$n = 10$, and $P = 0.4$

الحل:

(a) $P(r = 3) = 0.2150$

(b) $P(r \geq 8) = P(x = 8) + P(x = 9) + P(x = 10)$
$= 0.0106 + 0.0016 + 0.0001$
$= 0.0123$

(c) $P(2 \leq r \leq 5) = P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5)$
$= 0.1209 + 0.2150 + 0.2508 + 0.2007$
$= 0.7874$

3- توزيع بواسون

توزيع بواسون ويسمى أيضا توزيع الاحداث النادرة مثل حوادث المرور في شهر، الإصابة بمرض نادر. ويستخدم لتحديد احتمال وقوع عدد معين من النجاحات فى وحده الزمن ، وذلك عندما تكون الأحداث أو "النجاحات " مستقلة عن بعضها البعض وعندما يبقى متوسط عدد النجاحات ثابتاً لوحدة الزمن . عندئذ

$$P(X) = \frac{\lambda^x e^{-\lambda}}{X!}$$

حيث : **x** = العدد المعين من النجاحات.

P (x) = احتمال عدد **x** من النجاحات.

λ = (الحرف اليونانى لامدا)= متوسط عدد النجاحات فى وحدة الزمن .

e = أساس نظام اللوغاريتمات الطبيعى ، أو **2.71828**

وبمعلومية قيمة λ (القيمة المتوقعة أو متوسط وتباين توزيع بواسون) ، يمكن إيجاد $e^{-\lambda}$ من ملحق 2 ، والتعويض فى معادلة ، وإيجاد **p(x)**

يمكن استخدام توزيع بواسون كتقريب لتوزيع ذى الحدين عندما عدد المحاولات **n** كبيراً ، **P** أو **1−P** صغيرة (أحداث نادرة) . وكقاعدة عملية جيدة يستخدم توزيع بواسون عندما **n>30** و**np** أو **n(1−p)** أقل من **5** ، إذ عندما تكون **n** كبيرة فإن استخدام توزيع ذى الحدين يستغرق وقتاً طويلا كما أن جداول احتمالات ذى الحدين قد لا يكون متاحة للقيم الصغيرة للاحتمال **p** .

مثال: اذا كان متوسط الحوادث الأسبوعية في احدى الطرق في مدينة ما هو 3 حوادث احسب:-

1-احتمال أن يقع حادثين

2- احتمال ان يقع في أحد الاسابيع حادثان على الاقل.

3-احتمال ان يقع خمسة حوادث في اسبوعين.

4-متوسط عدد الحوادث على ذلك الطريق في السنة.

الحل:

لدينا: $\lambda= 3$

1- احتمال ان يقع حادثين:

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!} = \frac{3^2 e^{-3}}{2!} = 0.2240$$

2- احتمال ان يقع في أحد الاسابيع حادثان على الاقل.

$$P(X\geq x) = P(X\geq 2) = \sum_{x=2}^{\infty} P(X = x) = 1- P(X \leq x)$$

$$= 1-[P(X = 0)+ P(X = 1)] = 1-[\frac{\lambda^0 e^{-3}}{0!} + \frac{\lambda^1 e^{-3}}{1!}]$$

$$= 1- [e^{-3} + 3e^{-3}] = 0.8008$$

3- احتمال ان يقع خمسة حوادث في أسبوعين: اسبةعين يعني $\lambda= 6$

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!} = \frac{6^5 e^{-6}}{5!} = 0.1606$$

4- متوسط عدد الحوادث على ذلك الطريق في السنة

يوجد في السنة 48 أسبوعا اذن:

$$\lambda= 3x48 = 144$$

ويمكن استخدام توزيع بواسون كتقريب للتوزيع ذى الحدين عندما:

$n \geq 30$ و $np < 5$ أو $nq < 5$

مثلا: مقارنة بين توزيع ذي الحدين وبواسون عندما $\lambda = np, p = \frac{1}{100}, n = 100$

k	0	1	2	3	4	5
ذي الحدين	0.366	0.370	0.185	0.0610	0.0149	0.0029
بواسون	0.368	0.368	0.184	0.0613	0.0153	0.00307

ويمكن ملاحظة من خلال الجدول أعلاه التقارب الشديد للنتائج بين التوزيعين.

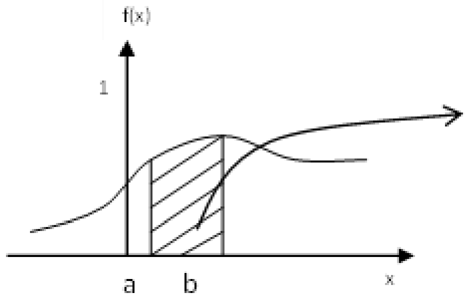
2-3-التوزيعات الاحتمالية المتصلة:

إذا كان المتغير العشوائي متصلاً مثل الدخل الشهري، كان التوزيع الاحتمالي له متصلاً. ولأن المتغير يأخذ عدداً لا نهائياً من القيم والذي لا يمكن حصره، فإنه يعبر عنها دائماً بدوال رياضية أو برسم منحنيات هذه الدوال أو التوزيعات.

ويعبر عن الاحتمالات بدلالة المساحات تحت هذه المنحنيات. وقد تم حساب الاحتمالات المختلفة لأهم التوزيعات المتصلة وتم وضعها في جداول إحصائية ثابتة يمكن استخدامها في أي وقت وفي أي مكان حيث توفر من الجهد والوقت في حساب هذه الاحتمالات، حيث يمكن الكشف في هذه الجداول في ثوان معدودة والحصول على الاحتمال المطلوب هو مجموعة القيم التي يمكن أن تأخذها المتغيرة ع المستمرة والاحتمالات الملحقة بها. نسمي توزيعا كهذا دالة الكثافة الاحتمالية أو دالة الاحتمال، وهي ممثلة بمنحنى متصل.

لاحظ أنه بما أن X تأخذ عددا لا متناهيا من القيم داخل أي مجال، فإن احتمال قيمة بعينها هو احتمال يؤول إلى الصفر.$P(X=x) \rightarrow 0$ لذلك فإن دالة الكثافة تستعمل لحساب احتمال مجال. ويكون ذلك بحساب المساحة تحت منحنى f(x) بين حدود المجال.

الشكل : التمثيل البياني لدالة الكثافة للمتغير العشوائي المتصل



$$P(a < x < b) = \int_a^b f(x)dx$$

.

خصائص دالة الكثافة الاحتمالية للمتغيرة العشوائية المستمرة

باستبدال إشارة التكامل بإشارة المجموع نجد أن شروط دالة

$$1)\ f(x) \geq 0$$

$$2)\int_{-\infty}^{+\infty} f(x)dx = 1$$

الكثافة الاحتمالية للمتغير العشوائي المتصل تكتب كما يلي :

من البديهي إذا أن منحنى دالة الكثافة لا يمكن أن ينزل أسفل محور الم ع، كما أن المساحة الإجمالية بين المنحنى والمحور الأفقي تساوي الواحد.

هذه الخصائص تفيدنا في حساب احتمالات بعض المجالات من خلال احتمالات مجالات أخرى.

مثال: أوجد قيمة الثابت C التي تحقق الشرطين الأول والثاني لدالة الكثافة الاحتمالية في الدالة التالية:

- ✓ أحسب احتمال أن تكون X تنتمي للمجال من 1 إلى 2.
- ✓ أحسب احتمال أن تكون X لا تنتمي للمجال من 1 إلى 2.

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{sinon} \end{cases}$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1 => \int_{-\infty}^{0} 0dx + \int_{0}^{3} Cx^2dx + \int_{3}^{+\infty} 0dx = 1 => C\left[\frac{x^3}{3}\right]_0^3 = 9C = 1 => C = 1/9$$

لكي تكون x دالة كثافة يجب أن يكون $C = 1/9$.

$$P(1 < x \leq 2) = \int_{1}^{2} f(x)dx = \int_{1}^{2} (1/9)x^2dx = \frac{1}{9}\left[\frac{x^3}{3}\right]_1^2 = \frac{1}{9}\left[\frac{8-1}{3}\right] = \frac{7}{27}$$

$$P(1 > x > 2) = 1 - P(1 < x < 2) = 1 - \frac{7}{27} = \frac{20}{27}$$

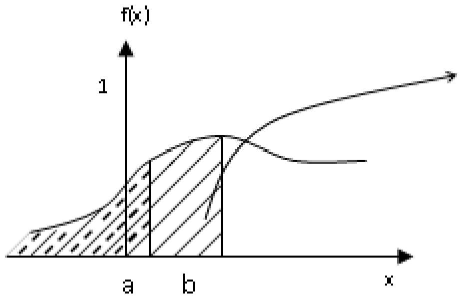
دالة التوزيع F(x) للمتغيرة العشوائية المستمرة

تعرف دالة التوزيع للمتغيرة المستمرة كما يلي:

$$\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du}$$

لدالة التوزيع أهمية أكبر بالنسبة للمتغيرة المستمرة. السبب في ذلك أننا نهتم، في حالة المتغيرة المستمرة، باحتمال مجال وليس باحتمال نقطة، ولحساب احتمال مجال من الأيسر التعويض في دالة التوزيع بدلا من حساب التكامل في كل مرة. يتضح ذلك من القاعدة التالية: بفرض a, b نقطتان من مجال تعريف X، بحيث b > a . لحساب احتمال أن تكون X تنتمي إلى المجال]a, b] :

$$\boxed{P(a < x \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)}$$



وهناك عدد مهم وشهير من التوزيعات الاحتمالية المتصلة أهمها:

2-3-1-التوزيع الطبيعى

هو أكثر التوزيعات استخداماً فى لعدة أسباب منها :

(1) له تطبيقات كثيرة ومتنوعة كتوزيع احتمالي لكثير من الظواهر الطبيعية مثل الذكاء، وأطوال الأشخاص وأوزانهم والظواهر الطبيعية (أو غير الطبيعية) الأخرى.

(2) ولاعتماد الكثير من الاختبارات الإحصائية والاستدلال الإحصائي بصفة عامة عليه

(3) وأيضا لاشتقاق عدد من التوزيعات المهمة – مثل توزيعات t، X^2، F من التوزيع الطبيعي.

أ– التوزيع الطبيعي : la loi normale ou D. de Laplace –Gausse : : يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتماليه لأنه يمثل كثيراً من الظواهر التي تقابلنا في الحياة العمليه مثل الأطوال و الأوزان والاعمار . ... و غيرها من الظواهر المتصلة .كما أن جميع التوزيعات الإحتمالية مثل توزيع ذو الحدين وتوزيع بواسون وغيرها من التوزيعات تؤول إلى التوزيع الطبيعى عندما يكون فضاء العينة كبير جدا. و تكون دالة الكثافة كالتالي :

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

- وهى دالة منحن له شكل الجرس متماثل حول المحور $X = \mu$ حيث:
- π عدد ثابت يساوى تقريبا 3.1416
- e عدد ثابت يساوى تقريبا 2.7183
- μ هى القيمة المتوقعة لـ X يمكن أن يكون أى عدد حقيقى $E(X) = \mu$

σ^2 هو التباين لـ X ويمكن أن يكون أى عدد حقيقى موجب $Var(X) = \sigma^2$

وهذه الدالة تحدد تماما متى علمت المعالم μ, σ

ونقول أن X له توزيع طبيعى بتوقع μ وتباين $\sigma 2$ وتكتب $X \sim N(\mu, \sigma 2)$

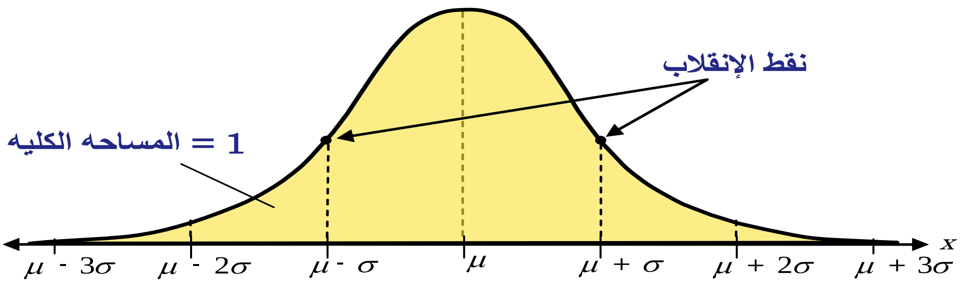
وبما ان المتغير مستمر و لحساب الاحتمالات يجب حساب التكامل التالي :

$$p(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

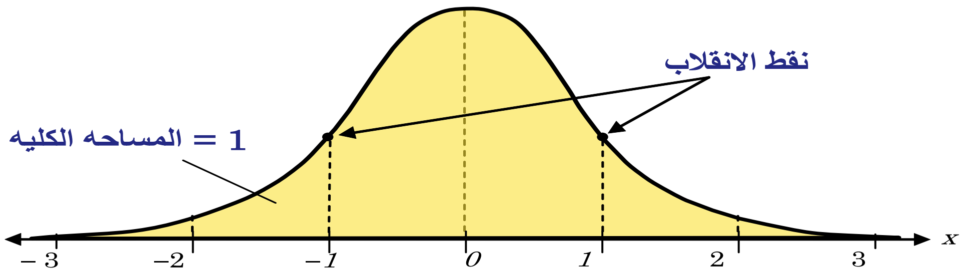
و هو امر صعب جدا فانه يقترح تحويل المنحنى المتماثل حول **$X = \mu$** إلى منحنى متماثل حول **$Z = 0$** حيث : $z=\left(\frac{x-\mu}{\sigma}\right)$

ليصبح المنحنى الإحتمالى للتوزيع الطبيعى القياسى (المعياري) بإستخدام المعادله

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}, \quad -\infty \le z \le \infty.$$



المنحنى الإحتمالى للتوزيع الطبيعى القياسى



خصائص المنحنى الإحتمالى للتوزيع الطبيعى القياسى

1. طرفا التوزيع تمتد من $-\infty$ إلى ∞
2. متماثل حول العمود الذى يمر بقمته أى عند $X=\mu=0$ لذلك فإن هذا العمود يجزئ المنحنى الطبيعى القياسى إلى قسمين متماثلين فى الشكل والمساحه.
3. يشبه الناقوس من حيث الشكل.
4. المنحنى له نقطتا إنقلاب عند $X=\pm 1$
5. المساحه الكليه تحت المنحنى تساوى الواحد الصحيح .
6. نظرا لتطابق جانبيه وتوسط تفرطحه فإن مقاييس النزعه المركزيه الثلاثه (الوسط الحسابى والوسيط والمنوال) تلتقى فى نقطه واحده (فى منتصف التوزيع) أى أن الوسط الحسابى = الوسيط = المنوال = 0
7. نلاحظ أن 99.7% من قيم المتغير تقع تقريبا فى الفتره $(-3, 3)$ أى أنه نادرا ما نجد قيمه من قيم X خارج هذه الفتره. كما أن 95.5% من قيم المتغير X تقع فى الفتره $(-2, 2)$ وأن 68 % منها تقع فى الفتره$(-1,1)$

ويمكن حساب إحتمال وقوع Z فى أى فتره تريدها وذلك بإيجاد قيمه تكامل داله الكثافه الإحتماليه فى هذه الفتره أى بإيجاد المساحه المحصوره تحت منحنى هذه الداله داخل هذه الفتره .

ولان متوسط التوزيع الطبيعى القياسى دائما يساوى صفر وتباينه يساوى واحد فقد تم عمل جدول يعطى إحتمالات وقوع المتغير Z فى فتره معينه .

مثال:

اذا كانت اعمار مجموعة من المسنين تتبع توزيعا طبيعيا متوسطه 34.8 سنة و بانحراف قدره 3.6 سنة . اخترنا شخصا ، ما هو احتمال ان يكون عمره :

1. اقل من 30 سنة ؟
2. اكبر من 40 سنة ؟
3. العمر الذي أقل منه **0.975** من الاشخاص
4. بين 28 و 38 سنة ؟

الحل:

قيمة معالم التوزيع الاحتمالي للعمر:

بفرض أن x متغير عشوائي يعبر عن الاعمار، وهو يتبع التوزيع الطبيعي، ومعالمه هي:

أ- المتوسط $E(x) = \mu = 34.8$ ب- التباين هو : $Var(x) = \sigma^2 = 12.96$

أي أن : $x \sim N(34.8, 12.96)$

1- نسبة الاشخاص الذين تقل أعمارهم عن 30 سنة هي: $P(x < 30)$

ويتبع الخطوات المذكورة سابقا في حساب الاحتمال كما يلي:

$$P(x < 30) = p\left(Z < \frac{X-\mu}{\sigma} \right)$$

$$P(x < 30) = p\left(Z < \frac{30-34.8}{3.6} \right)$$

$$= p\left(Z < \frac{30-34.8}{3.6} \right) = p(z < -1.33$$

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

$$P(x < 30) = P(z < -1.33) = 0.09176$$

2-نسبة الاشخاص الذين تزيد أعمارهم عن 40 سنة هي$P(x>40)$:

$$P(x > 40) = p\left(Z > \frac{X-\mu}{\sigma} \right)$$

$$P(x > 40) = p\left(Z > \frac{40-34.8}{3.6} \right)$$

$$= p\left(Z > \frac{30-34.8}{3.6} \right) = p(z > 2.78$$

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

$$P(x > 40) = P(z > 2.78) = 0.99632$$

3-العمر الذي أقل منه 0.975 من الاشـــــخاص: في هذه الحالة يبحث عن قيمة المتغير (x) الذي أقل منه 0.975 ، بفرض أن هذا المتغير هو (x_1) ، فإن :

$$P(x < x_1) = p\left(z < \frac{x_1 - 34.8}{3.6}\right) = 0.975$$

بالكشـف بطريقة عكسـية ، حيث نبحث عن المسـاحة 0.9750 نجدها تقع عند تقاطع الصـف 1.9 ، والعمود $.06$. أي أن قيمة $z = 1.96$ ، ويكون :

$$1.96 = \frac{x_1 - 34.8}{3.6} \quad \Rightarrow x_1 = 3.6(1.96) + 34.8 = 41.86$$

إذا العمر هو $.41.86$ سنة.

4-نسبة الاشخاص الذين تتراوح أعمارهم بين 28 و 38. سنة

$$\boldsymbol{Z\ 1 = \frac{X-\mu}{\sigma} = \frac{28-34.8}{3.6} = -1.89 \Rightarrow p = 0.81327}$$

$$\boldsymbol{Z\ 1 = \frac{X-\mu}{\sigma} = \frac{38-34.8}{3.6} = 0.89 \Rightarrow p = 0.02938}$$

أي نسبة الاشخاص الذين تتراوح أعمارهم بين 28 و 38. سنة تكون محصورة بين 0.02938 و 0.81327.