مبادئ الاحصاء والاحتمالات

Principle of Statistic and Probability

MS 304

محاضرة رقم 11

المتغيرات العشوائية والتوزيعات الاحتمالية

Random Variables and Probability Distributions

المتغير العشوائي Random Variable:

المتغير العشوائي هو الذي يأخذ قيما حقيقية مختلفة تعبر عن نتائج فراغ العينة، ومن ثم مجـــال هذا المتغير، يشمل كل القيم الممكنة له، ويكون لكل قيمة من القيم التي يأخذها المتغير احتمال معين، وينقسم المتغير العشوائي إلى قسمين هما:

1- المتغيرات العشوائية المنفصلة Discrete Random Variables

2- المتغيرات العشوائية المتصلة(المستمرة) Continuous Random Variables

المتغيرات العشوائية المنفصلة

المتغير العشوائي المنفصل هو الذي يأخذ قيم بينية، ومتباعدة، ويرمز للمتغير العشوائي بشكل عام بحرف من الحروف الأبجدية الكبيرة $X, Y, Z, \ldots$ ويرمز للقيم التي يأخذها المتغير بالحروف الأبجدية الصغيرة، $x, y, z, \ldots$ ومن أمثلة هذه المتغيرات:

1- عدد الأولاد الذكور في الأسرة المكونة من أربع أولاد X، $X:\{x=0,1,2,3,4\}$.

2- عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق Y، $Y:\{y=0,1,2,3,\ldots\}$.

3- عدد مرات استخدام نوع معين من الأسمدة خلال الدورة الزراعية.

4- عدد الوحدات التالفة من إنتاج مزرعة معينة تنتج 200 وحدة كل موسم.

5- عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر.

التوزيع الاحتمالي للمتغير العشوائي المنفصل

التوزيع الاحتمالي، هو الذي يبين احتمالات حدوث القيم التي يمكن يأخذها المتغير، والتي تـــرتبط باحتمالات النتائج الممكنة في فراغ العينة، وبمعنى آخر هو التكراري النسبي للقيم التي يمكن أن يأخذها المتغير.

فإذا كان المتغير العشوائي المنفـــصل X يأخـــذ القـــيم، $X:\{x=x_1,x_2,...,x_n\}$، وكـــان $P(X=x_i)=f(x_i)$ هو احتمال أن المتغير العشوائي يأخذ القيمة x_i، فإنه، يمكـــن تكـــوين جدول التوزيع الاحتمالي للمتغير العشوائي X، وهو جدول مكون من عمودين، الأول به القـــيم الممكنة للمـــتغير $X:\{x=x_1,x_2,...,x_n\}$، والثـــاني بـــه القـــيم الاحتماليـــة لهـــذا المـــتغير $P(X=x_i)=f(x_i)$، أي أن:

جدول التوزيع الاحتمالي للمتغير العشوائي المنفصل

x_i	$f(x_i)$
x_1	$f(x_1)$
x_2	$f(x_2)$
$\vdots$	$\vdots$
x_n	$f(x_n)$
Σ	1

وتسمى الدالة $f(x_i)$ بدالة الاحتمال، ومن خصائص هذه الدالة ما يلي:

$$1- \quad 0 < f(x_i) < 1$$

$$2- \quad \sum f(x_i) = 1$$

التوزيعات الاحتمالية المنفصلة الخاصة

في كثير من النواحي التطبيقية، تتبع بعض الظواهر توزيعات احتمالية خاصة، وهي التوزيعات التي يمكن حساب احتمالات قيم المتغير عن طريق معادلة رياضية، تسمى بدالة الاحتمال $f(x)$ ، وهذه المعادلة لها معالم معينة، تسمى بمعالم المجتمع الذي ينسب له هذا التوزيع، وهذه المعالم ما هي إلا حقائق ثابتة مجهولة، وهي الأساس في حساب القيم الاحتمالية للتوزيع الاحتمالي للمجتمع محل الدراسة.

ومن أهم التوزيعات التي سيتم دراستها في هذا المقرر، توزيع ثنائي الحدين:

التوزيع ثنائي الحدين The Binomial Distribution

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح، والأخرى تسمى بحالة الفشل، ومن أمثلة ذلك:

- عند إعطاء مريض نوع معين من الأدوية، لها نتيجتان: (استجابة للدواء، أو عدم استجابة)
- عند فحص عبوة بداخلها نوع معين من الفاكهة، لها نتيجتان (الوحدة إما أن تكون سليمة، أو تكون معيبة)
- عند إلقاء قطعة عملة، لها نتيجتان (ظهور الوجه الذي يحمل الصورة، أو الوجه الذي يحمل الكتابة)
- نتيجة الطالب في الاختبار (نجاح، رسوب)

شكل التوزيع الاحتمالي ثنائي الحدين

إذا كررت محاولة n من المرات، بحيث أن كل محاولة لها نتيجتان فقط متنافيتان هما:

- النتيجة محل الاهتمام " حالة نجاح " وتتم باحتمال ثابت في كل محاولة هو p
- النتيجة الأخرى " حالة فشل " وتتم باحتمال ثابت أيضا هو $q = 1 - p$

وبافتراض أن هذه المحاولات مستقلة، بمعنى أن نتيجة كل محاولة ليس لها علاقة بنتيجة المحاولة الأخرى، وإذا كان المتغير العشوائي X يعبر عن عدد حالات النجاح "عدد النتائج محل الاهتمام" في الـ n محاولة، فإن مدي المتغير العشوائي X والذي يعبر عن عدد حالات النجاح هو: $X : \{x = 0,1,2,\ldots,n\}$ ، ومن ثم يحسب الاحتمال $P(X = x) = f(x)$ بتطبيق المعادلة التالية:

$$\boxed{f(x) = \binom{n}{x} p^x q^{n-x} \ , \ x = 0,1,2,\ldots n}$$

حيث أن $\binom{n}{x}$ هي عدد طرق اختيار x من n مع إهمال الترتيب، وتحسب كما يلي:

$$\boxed{\binom{n}{x} = \frac{n(n-1)(n-2)\ldots(n-x+1)}{x(x-1)(x-2)\ldots 3\times 2\times 1}}$$

$$\binom{7}{3} = \frac{7\times 6\times 5}{3\times 2\times 1} = 35 \quad = \binom{7}{4}$$

$$\binom{7}{0} = \binom{7}{7} = 1$$

مثـــــــال

إذا كان من المعلوم أن نسبة الشفاء من مرض معين باستخدام نوع معين من العقاقير الطبية هو 0.60، إذا تناول هذا العقار 5 مصابين بهذا المرض. إذا عرف المتغير العشوائي X بأنه عدد الـــذين المستجيبين (حالات الشفاء) لهذا العقار.

المطلوب:

أ- ما هو نوع المتغير؟

ب-اكتب شكل دالة الاحتمال $f(x)$ لهذا المتغير.

ت-احسب الاحتمالات التالية:

- ما احتمال استجابة 3 مرضى لهذا العقار؟
- ما هو احتمال استجابة مريض واحد على الأقل؟
- ما هو احتمال استجابة 2 مرضى على الأكثر؟

الحـــــل:

أ- عدد حالات الاستجابة X متغير كمي منفصل ، ومدى هذا المتغير في هذه الحالة هـــو:

$: X : \{x = 0,1,2,3,4,5\}$

ب- شكل دالة الاحتمال: $n = 5$ ، $p = 0.60$ ، $q = 1 - p = 0.40$ إذا:

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

$$= \binom{5}{x}(0.6)^x (0.4)^{5-x} \quad , \quad x = 0,1,2,3,4,5$$

ت- حساب الاحتمالات:

- حساب احتمال استجابة 3 مرضى لهذا الدواء: $P(x = 3) = f(3)$

$$f(3) = \binom{5}{3}(0.6)^3 (0.4)^{5-3} = \frac{5 \times 4 \times 3}{3 \times 2 \times 1} \times 0.216 \times 0.16 = 10 \times 0.03456$$

$$= 0.3456$$

- حساب احتمال استجابة مريض واحد على الأقل: $P(x \geq 1)$

$$P(x \geq 1) = f(1) + f(2) + f(3) + f(4) + f(5) = 1 - f(0)$$
$$= 1 - \left[\binom{5}{0}(0.6)^0(0.4)^5\right] = 1 - 1 \times 1 \times 0.01024 = 0.98976$$

- حساب احتمال استجابة 2 مرضى على الأكثر: $P(x \leq 2)$:

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

$$= \binom{5}{2}(0.6)^2(0.4)^3 + \binom{5}{1}(0.6)^1(0.4)^4 + \binom{5}{0}(0.6)^0(0.4)^5$$
$$= \frac{5 \times 4}{2 \times 1}(0.36)(0.064) + \frac{5}{1}(0.6)(0.0256) + 1(1)(0.01024$$
$$= 0.2304 + 0.0768 + 0.01024 = 0.31744$$

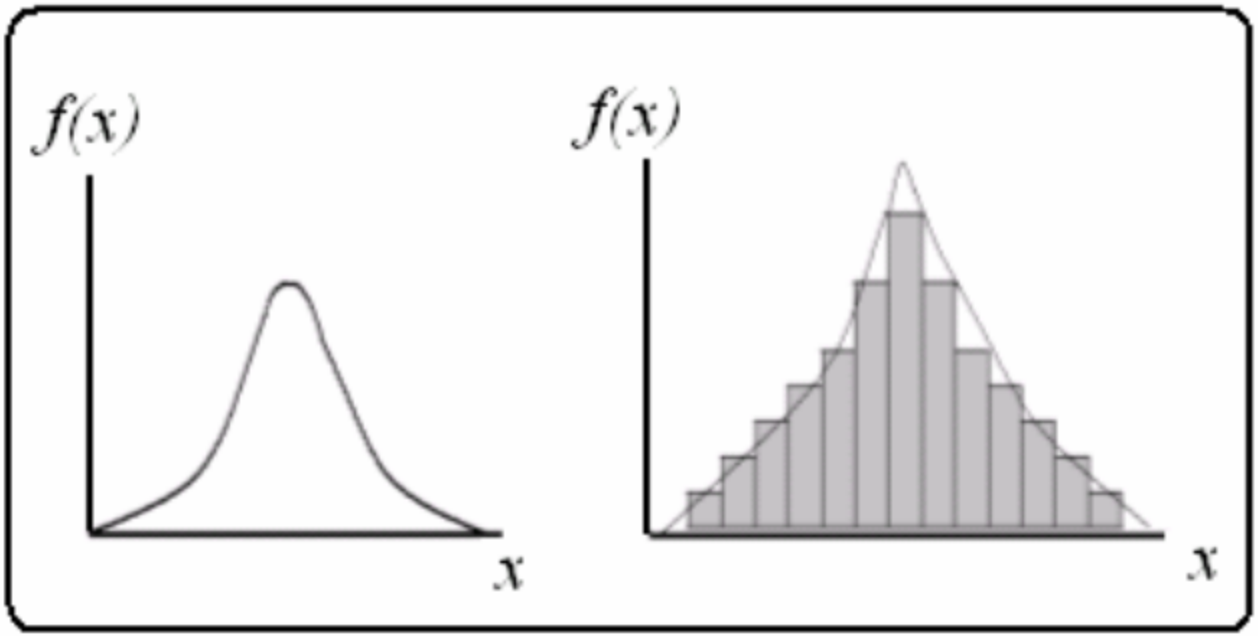
المتغيرات العشوائية المستمرة Continuous Random Variables

المتغير العشوائي المستمر، هو الذي يأخذ قيما متصلة، ويأخذ عدد لانهائي من القيم الممكنة له داخل مجاله، فإذا كان X متغير عشوائي مستمر، ويقع في المدى (a,b)، أي أن: $\{X = x : a < x < b\}$، فإن للمتغير X عدد لانهائي من القيم تقع بين الحدين الأدنى والأعلى (a,b)، ومن الأمثلة على المتغيرات الكمية المستمرة ما يلي:

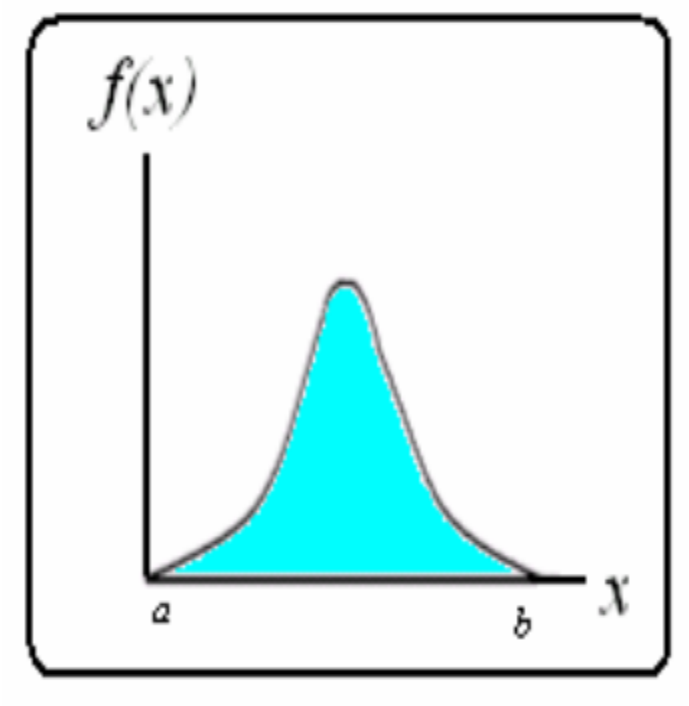
- كمية الألبان التي تنتجها البقرة في اليوم باللتر: $\{X = x : 10 < x < 40\}$
- المساحة المترعة بالأعلاف في المملكة بالألف هكتار $\{X = x : 1000 < x < 15000\}$
- فترة صلاحية حفظ الدجاج المبرد بالأيام، $\{X = x : 1 < x < 5\}$
- وزن الجسم بالكيلوجرام للأعمار من $(30\text{-}40)$، $\{X = x : 55 < x < 80\}$

التوزيع الاحتمالي للمتغير المستمر Continuous Probability

عند تمثيل بيانات المتغير الكمي المستمر في شكل مدرج تكراري النسبي، نجد أن شكل هذا المدرج هو أقرب وصف لمنحنى التوزيع الاحتمالي للمتغير المستمر، وكلما ضاقت الفترات بين مراكز الفئات، يمكن الحصول على رسم دقيق للمنحنى الخاص بدالة احتمال المتغير المستمر، كما هو مبين بالشكل التالي:


f(x)
x
f(x)
x

والمساحة أسفل المنحنى تعبر عن مجموع الاحتمالات الكلية، ولذا تساوي هذه المساحة الواحد الصحيح، وتسمى الدالة $f(x)$ بدالة كثافة الاحتمال Probability Distribution Function$(p.d.f)$ ، وبفرض المتغير العشوائي المستمر يقع في المدى: $X = \{x : a \pi x \pi b\}$ ، وأن منحنى هذه الدالة يأخذ الصورة التالية:



فإن من خصائص دالة كثافة الاحتمال $f(x)$ ما يلي:

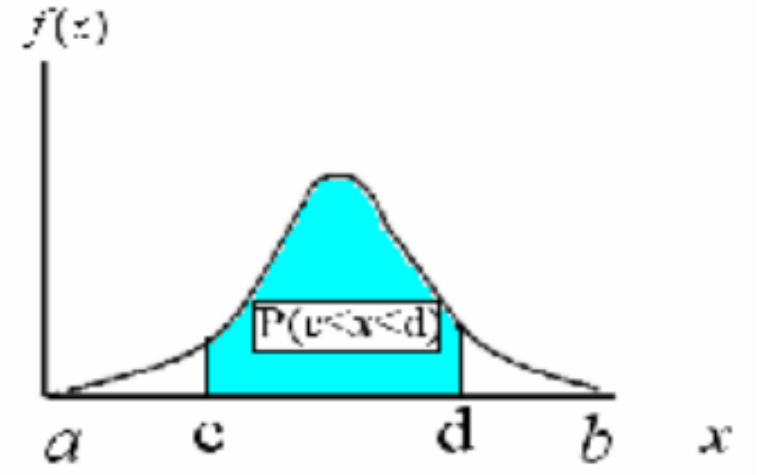
1- الدالة $f(x)$ موجبة داخل المدى (a,b) أي أن: $f(x) \phi 0$ ، $x \in (a,b)$

2- التكامل على حدود المتغير من الحد الأدنى a حتى الحد الأعلى b يعبر عن مجموع الاحتمالات الكلية، لذا يساوي الواحد الصحيح ، أي أن:

$$\boxed{\int_{x=a}^{x=b} f(x)\ dx = 1}$$

حيث أن الشكل الرياضي أعلاه يسمى بالتكامل المحدد من $x = a$ حتى $x = b$ ، وهذا يعني إيجاد المساحة أسفل المنحني بين (a,b) .

3- لحساب احتمال أن المتغير العشوائي المستمر يقع في المدى (d,c) أي حساب الاحتمال $p(c < x < d)$ ، يجب حساب المساحة أسفل المنحني من $x = c$ حتى $x = d$ كما هي مبينة في الشكل البياني التالي:



ويتم ذلك بإيجاد التكامل المحدد في هذا المدى، كما يلي:

$$p(c < x < d) = \int_{x=c}^{x=d} f(x)\ dx = [g(x)]_c^d = g(d) - g(c)$$

4- في المتغير المستمر، يكون الاحتمال $p(x = value)$ مساويا للصفر، أي أن:

$$p(x = value) = 0$$