الفصل الرابع

التوزيع المعتدل والتوزيع المعتدل اللوغاريتمى

THE NORMAL DISTRIBUTION AND THE LOGNORMAL DISTRIBUTION

(أولا) التوزيع المعتدل

التوزيع المعتدل هو أهم توزيعات الاحتمال القياسية وأكثرها استخداما ، إذ يؤخذ كنموذج لكثير من المتغيرات المتصلة ، منها أطوال بتلات بعض النباتات – أطوال أجنحة الذباب المنزلى – أطوال وموازين الأطفال عند الولادة – مقادير الدسم فى الزبد الناتج من بعض أنواع الأبقار ... فضلا عن أنه يلعب دورا كبيرا فى بناء نظرية الاحتمال والنظرية الإحصائية .

وقد سمى هذا التوزيع بالتوزيع المعتدل (أو المعتاد أو الطبيعى) لأنه كان يظن فيما مضى أن أية بيانات عن ظواهر الحياة ينبغى أن تمتثل لهذا التوزيع وإلا كانت هذه البيانات مشكوكا فيها . إلا أنه ثبت الآن أن الأمر ليس كذلك فهناك كثير من الظواهر ذات توزيعات تختلف عن التوزيع المعتدل . كما يسمى التوزيع بتوزيع جاوس اعترافا بفضل العالم الألمانى كارل فردريك جاوس (١٧٧٧ – ١٨٥٥) الذى استنبط التوزيع رياضيا كتوزيع احتمال أخطاء القياس وكان لذلك يسميه

بالقانون الطبيعى للأخطاء normal law of errors ، ويسمى أيضا بتوزيع جاوس - لابلاس اعترافا بجميل العالم الفرنسى بيير سيمون لابلاس (١٧٤٩ - ١٨٢٧) الذى كان أول من اكتشفه وأثبت صلاحيته كنموذج لكثير من الظواهر .

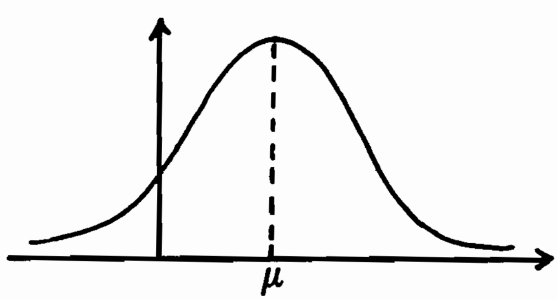
ويعرّف هذا التوزيع بواسطة دالة كثافة الاحتمال المعطاة بالقاعدة :

$$د(س) = \frac{1}{\sqrt{2ط\,\sigma^{2}}}\; هـ^{-\frac{1}{2\sigma^{2}}(س-\mu)^{2}} \qquad حيث \quad \infty > س > -\infty ،\ \sigma > 0 \qquad (1)$$

ولهذه الدالة بارامتران هما μ ، σ ، تعبر μ عن الوسط الحسابى وتعبر σ عن الانحراف المعيارى للتوزيع . ويتحدد التوزيع تماما إذا علمت قيمتا هذين الدليلين ، ولذا نرمز له بالرمز مع (μ، σ) . والمتغير العشوائى س الذى له هذه الدالة يسمى بالمتغير المعتدل .

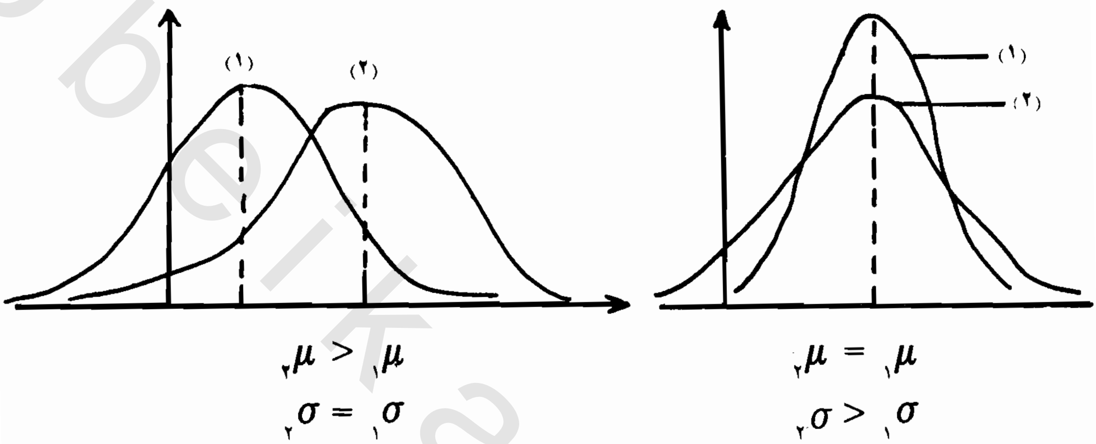
(٤ - ١) بعض خواص التوزيع

(أ) المنحنى الذى يمثل الدالة (١) يسمى بالمنحنى المعتدل وهو منحنى ذو قمة واحدة ومتماثل حول الخط س = μ . ومساحة المنطقة الواقعة بين هذا المنحنى ومحور السينات تساوى الواحد الصحيح كما هو الحال لمنحنى أى توزيع احتمال . ومن التماثل نجد أن الخط س = μ يقسم الشكل إلى منطقتين متساويتى المساحة ومساحة كل منهما تساوى $\frac{1}{2}$. انظر الشكل (٤ - ١) .



الشكل (٤ - ١) : منحنى التوزيع المعتدل

(ب) لكل من البارامترين μ ، σ عدد غير منتهى من القيم ، ولذلك هناك عدد غير منتهى من التوزيعات المعتدلة . هذا مع ملاحظة أن μ هو بارامتر موضع location parameter أما σ فهو بارامتر شكل shape parameter كما يتبين من الشكل (٤ — ٢) الآتى :



الشكل (٤ - ٢)

(جـ) معامل الالتواء = صفر ، معامل التفرطح = ٣

(د) إذا كانت س ترمز إلى قيم متغير معتدل س وسطه الحسابي μ وانحرافه المعيارى σ فإن الصيغة :

$$ع = \frac{س - \mu}{\sigma} \qquad (٢)$$

تحول المتغير س إلى متغير ع له توزيع معتدل وسطه الحسابي صفر وانحرافه المعيارى الواحد الصحيح ، ويسمى هذا المتغير بالمتغير المعتدل المعيارى وسنرمز له بالرمز : مع (٠ ، ١) .

(هـ) إذا رسم توزيع التكرارات النسبية المتجمعة لمتغير معتدل على ورق الرسم البياني ذى التقسيم الخطى (العادى) فإن هذا المنحني يتخذ الشكل S غير أن

هناك ورق يسمى ورق تقسيم الاحتمالات المعتدلة normal probability graph paper إذا رسم عليه التوزيع نتج خط مستقيم ، ويستخدم هذا الورق للكشف عن اعتدالية المجتمعات كما سنرى في البند (٤ – ٤) الآتي .

(٤ – ٢) جداول المساحات أسفل المنحني المعتدل المعيارى :

نظراً لأهمية معرفة احتمالات المتغيرات المعتدلة وكثرة الحاجة إليها فقد حسبت قيم هذه الاحتمالات لمختلف فترات المتغير المعتدل المعيارى ع ووضعت في جداول تعرف بجداول المساحات أسفل المنحني المعتدل المعيارى ، وهذه الجداول تأخذ صوراً متعددة تؤدى إلى نفس النتائج ومنها الصورة الواردة بالجدول (٦) بملحق هذا الكتاب . وهذا الجدول يعطى المساحة تحت المنحني المعتدل المعيارى وفوق الفترة بين الصفر وبعض أعداد موجبة أ متوقعة للمتغير ع ، وهذه هى المساحة المظللة بالشكل (٤ – ٣) وهى تعبر عن الاحتمال ل (أ ≥ ع > ٠) (٣)

الشكل (٤ – ٣)

أى عن احتمال وقوع قيم المتغير ع بين العددين . ، أ أى في الفترة (. ، أ) . ويلاحظ من تماثل المنحني أن هذا الاحتمال هو أيضاً احتمال وقوع قيم المتغير ع بين العددين – أ ، . أى : ل (٠ ≥ ع > – أ) (٤)

وذلك لتساوى مساحتي المنطقتين المناظرتين .

في هذا الجدول يعبر الهامش الرأسي (الذى على اليسار أو على اليمين) عن

قيم أ إلى خانة عشرية واحدة ، ويعبر الهامش الأفقي (الذى في أعلى الجدول) عن الخانة العشرية الثانية أى خانة الجزء من مائة . أما الأعداد التي في قلب الجدول فهى احتمالات وقوع المتغير ع في الفترة (٠ ، أ) .

مثال (٤ − ١) :

إذا علم أن توزيع درجات الطلاب في مادة ما هو توزيع معتدل وسطه الحسابي μ = ٦٠ وانحرافه المعيارى σ = ١٦ . فأوجد باستخدام جدول المساحات الاحتمالات الآتية :

أولا : ل (٦٠ < س ≤ ٨٠)   **ثانياً** : ل (س > ٨٠)

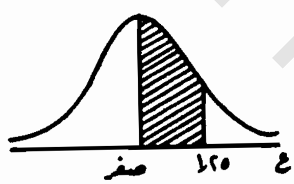
ثالثاً : **ل (س ≤ ٨٠)**   **رابعاً** : ل (٤٠ < س ≤ ٧٦)

خامساً : ل (٧٥ < س ≤ ٨٥)

الحل :

لإيجاد الاحتمالات المطلوبة من الجدول يجب أن نحول المتغير المعتدل س إلى المتغير المعتدل المعيارى ع بواسطة الصيغة (٢) وهى هنا

$$ع = \frac{س - ٦٠}{١٦}$$



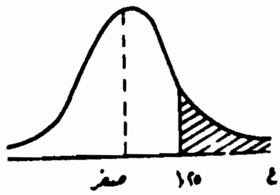
(أولاً) بوضع س = ٦٠ نجد أن ع = $\frac{٦٠ - ٦٠}{١٦}$ = صفر

وبوضع س = ٨٠ نجد أن ع = $\frac{٨٠ - ٦٠}{١٦}$ = ١,٢٥

إذن ل (٦٠<س≤٨٠) = ل (٠< ع ≤ ١,٢٥)

= ٠,٣٩٤٤

وذلك من الجدول مباشرة عند العدد ١,٢ الذى في الهامش الرأسي وتحت العدد ٥ الذى في الهامش الأفقي . وهذه النتيجة تعني أن حوالى ٣٩٪ من الطلاب تقع درجاتهم بين ٦٠ ، ٨٠ .



(ثانيا) بوضع س = ٨٠ نجد أن ع = ١,٢٥

∴ (س > ٨٠) = ل (ع > ١,٢٥)

= مساحة المنطقة التي على يمين العدد ١,٢٥

= ٠,٥٠٠٠ − ٠,٣٩٤٤

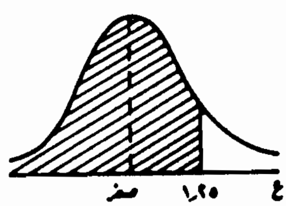
= ٠,١٠٥٦

أى أن حوالى ١١٪ من الطلاب تزيد درجاتهم عن ٨٠ .

(ثالثاً) ل (س ≤ ٨٠) = ل (ع ≤ ١,٢٥)

= مساحة المنطقة التي على يسار العدد ١,٢٥

= ٠,٣٩٤٤ + ٠,٥٠٠٠ = ٠,٨٩٤٤

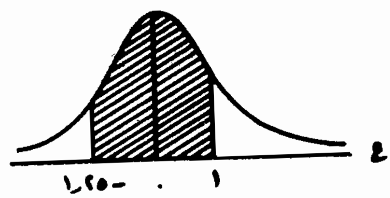


أى أن حوالى ٨٩٪ من الطلاب درجاتهم تساوى أو تقل عن ٨٠ .

يلاحظ أن جواب أى من الاسئلة الثلاثة السابقة يمكن الحصول عليه من جوابى السؤالين الآخرين .

(رابعاً) بوضع س = ٤٠ نجد أن ع = $\frac{٤٠ - ٦٠}{١٦}$ = −١,٢٥

بوضع س = ٧٦ نجد أن ع = $\frac{٧٦ - ٦٠}{١٦}$ = ١



إذن ل (٧٦ ≥ س > ٤٠) = ل (١ ≥ ع > − ١,٢٥)

= مساحة المنطقة فوق الفترة (٠ ، ١) + مساحة المنطقة فوق الفترة (− ١,٢٥ ، ٠) = ٠,٣٤١٣ + ٠,٣٩٤٤ = ٠,٧٣٥٧

(خامساً) س = ٧٥ تعطى ع = ٠,٩٣٧٥ = ٠,٩٤ تقريباً

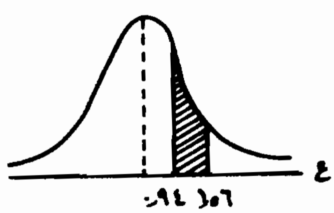
، س = ٨٥ تعطى ع = ١,٥٦٢٥ = ١,٥٦

إذن ل (٨٥ ≥ س > ٧٥) = ل (١,٥٦ ≥ ع > ٠,٩٤)

= المساحة فوق الفترة (٠ ، ١,٥٦)

− المساحة فوق الفترة (٠ ، ٠,٩٤)

= ٠,٤٤٠٦ − ٠,٣٢٦٤ = ٠,١١٤٢



مثال (٤ − ٢) : مثال مشهور

للمتغير المعتدل س الذى وسطه الحسابي μ وانحرافه المعيارى σ اثبت أن :

(أولا) احتمال وقوع قيم المتغير بين العددين $\mu \pm \sigma$ هو ٠,٦٨٢٦

(ثانيا) احتمال وقوع قيم المتغير بين العددين $\mu \pm ٢\sigma$ هو ٠,٩٥٤٤

(ثالثا) احتمال وقوع قيم المتغير بين العددين $\mu \pm ٣\sigma$ هو ٠,٩٩٧٤

الحل :

لكى نستخدم الجدول نحول المتغير س إلى المتغير ع بواسطة التعويض :

$$ع = (س - \mu) / \sigma$$

أولاً : بوضع : س = $\mu - \sigma$ نجد أن ع = $\frac{(\mu - \sigma) - \mu}{\sigma}$ = −١

وبوضع : س = $\mu + \sigma$ نجد أن ع = $\frac{(\mu + \sigma) - \mu}{\sigma}$ = ١

∴ ل $(\mu + \sigma \geqslant$ س $> \mu - \sigma)$ = ل(١ ⩾ ع > − ١)

= ٢ ل (١ ⩾ ع > ٠) من التماثل

= ٢ × ٠,٣٤١٣ =٠,٦٨٢٦

ثانياً : بالمثل نجد أن :

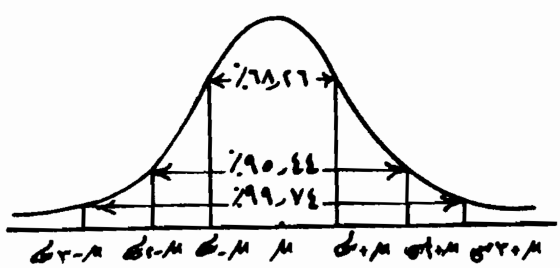
ل $(\mu + ٢\sigma \geqslant$ س $> \mu - ٢\sigma)$ = ل (٢ ⩾ ع > − ٢)

= ٢ ل (٢ ⩾ ع > ٠) = ٢ × ٠,٤٧٧٢ = ٠,٩٥٤٤

ثالثاً : ل $(\mu + ٣\sigma \geqslant$ س $> \mu - ٣\sigma)$ = ل (٣ ⩾ ع > − ٣)

= ٢ ل (٣ ⩾ ع > ٠) = ٢ × ٠,٤٩٨٧ = ٠,٩٩٧٤

وتصور هذه النتائج بيانياً كما في الشكل (٤ − ٤) الآتي :



الشكل (٤ − ٤) المساحات أسفل المنحني المعتدل

مثال (٤ – ٣) : مثال مشهور

بنفس الطريقة نثبت العلاقات الهامة الآتية التي سنحتاج إليها في كثير من التطبيقات الإحصائية ، انظر السؤال (١ – حـ) من تمارين (٤) .

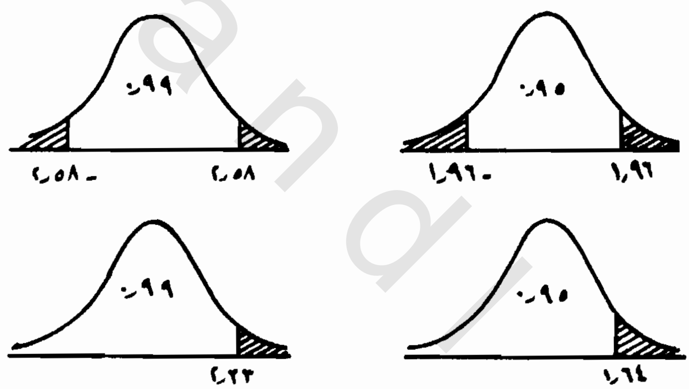
(١) ل (١,٩٦ ≥ ع > – ١,٩٦) = ٠,٩٥

(٢) ل (٢,٥٨ ≥ ع > – ٢,٥٨) = ٠,٩٩

(٣) ل (ع > ١,٦٤) = ٠,٠٥

(٤) ل (ع > ٢,٣٣) = ٠,٠١

وتمثل هذه الاحتمالات كما في الشكل (٤ – ٥) الآتي :



الشكل (٤ – ٥) بعض القيم الحرجة للمتغير المعتدل المعيارى

(٤ – ٣) الكشف عن الاعتدالية :

في كثير من الأحيان يبني التحليل الإحصائي لبيانات ناتجة من عينة على أساس افتراض أن المجتمع الذى سحبت منه العينة معتدل ، ولذلك ينبغى أن نتحقق من توفر هذا الافتراض قبل إجراء مثل هذا التحليل .

نفرض أن لدينا توزيعاً تكرارياً ذا فئات لعينة عشوائية وسطها الحسابي س̄ وانحرافها المعيارى ع ونريد اختبار ما إذا كانت هذه العينة مأخوذة من مجتمع معتدل .

نتصور مجتمعاً معتدلا له نفس الوسط الحسابي والانحراف المعيارى للتوزيع التكرارى المعلوم أى نأخذ μ = س̄ ، σ = ع . إذا كانت العينة مأخوذة من هذا المجتمع فإن احتمال وقوع المتغير المعتدل في فئة مساوية لأى من الفئات التي ينقسم إليها التوزيع التكرارى لا يجب أن يختلف كثيراً عن التكرار النسبي المشاهد في هذه الفئة . وعلى ذلك نقوم بحساب احتمالات وقوع المتغير المعتدل في جميع فئات التوزيع التكرارى مستعينين في ذلك بجدول المساحات . بضرب هذه الاحتمالات (التكرارات النسبية) في حجم العينة نحصل على ما يسمى بالتكرارات النظرية أو المتوقعة المناظرة للتكرارات المشاهدة في العينة . نقارن بين التكرارات المشاهدة والتكرارات المتوقعة لها فإذا كانت قريبة من بعضها بدرجة معقولة أى كانت هناك مطابقة حسنة بين التوزيع التكرارى المشاهد والتوزيع التكرارى النظرى ، جاز لنا أن نعتبر أن المجتمع معتدل ، وإلا فهو ليس كذلك .

إن عملية إيجاد توزيع تكرارى نظرى بالطريقة المذكورة تسمى بعملية **توفيق توزيع معتدل لتوزيع تكرارى معلوم** . وقد سبق أن مرت بنا فكرة التوفيق هذه في حالة كل من توزيع ذى الحدين وتوزيع بواسون . وكما سبق القول ، يعتمد الحكم الموضوعى على حسن المطابقة أو سوئها على أحد الاختبارات الإحصائية مثل اختبار χ^2 الذى سندرسه فيما بعد .

مثال (٤ - ٤) :

وفق التوزيع المعتدل للتوزيع التكرارى الآتي ، وإذا كان هذا التوزيع لعينة عشوائية فاختبر ما إذا كان من الممكن اعتبار أن المجتمع الذى سحبت منه هو مجتمع معتدل .

الفئة	٣١٥-	٣٢٥-	٣٣٥-	٣٤٥-	٣٥٥-	٣٦٥-	٣٧٥-	٣٨٥-	٣٩٥-	٤٠٥-
التكرار	٦	٦	١١	١٤	١٦	١٥	٨	١٠	٨	٦

الحل :

يتطلب توفيق التوزيع المعتدل أن نوجد الوسط الحسابي والانحراف المعيارى للتوزيع التكرارى المعطى لاستخدامهما في تقدير الوسط الحسابي والانحراف المعيارى للمجتمع . كالمعتاد نجد أن :

الوسط الحسابى $\bar{\text{س}} = \frac{1}{\text{ن}} \sum \text{ك}_{\text{ر}} \text{س}_{\text{ر}} = \frac{1}{100}(6\times320+6\times330+\dots+6\times410) = \frac{36470}{100}$

$= 364.7$

التباين $= \text{ع}^2 = \frac{1}{\text{ن}-1}\left(\sum \text{ك}_{\text{ر}} \text{س}_{\text{ر}}^2 - \frac{(\sum \text{ك}_{\text{ر}} \text{س}_{\text{ر}})^2}{\text{ن}}\right)$

$= \frac{1}{99}\left(6\times320^2+6\times330^2+\dots+6\times410^2 - \frac{36470^2}{100}\right)$

$= 712.89$

الانحراف المعيارى $= 26.7$

نريد أن نختبر أن المجتمع الذى سحبت منه العينة هو مجتمع معتدل وسطه الحسابي ٣٦٤,٧ وانحرافه المعيارى ٢٦,٧ . نحسب التكرارات المتوقعة كما في الجدول (٤ – ١) الآتي :

الجدول (٤ - ١)

توفيق توزيع معتدل للتوزيع التكراري في المثال (٤ - ٤)

الفئات بقيم س	الفئات بقيم ع	التكرارات النسبية المتوقعة $ل_ر$	التكرارات المتوقعة $ع_ر = ل_ر \times ١٠٠$	التكرارات المشاهدة $ك_ر$
-∞-٣٢٥	-∞-(-١,٤٩)	٠,٠٦٨١	٦,٨١	٦
٣٢٥-٣٣٥	(-١,٤٩)-(-١,١١)	٠,٠٦٥٤	٦,٥٤	٦
٣٣٥-٣٤٥	(-١,١١)-(-٠,٧٤)	٠,٠٩٦١	٩,٦١	١١
٣٤٥-٣٥٥	(-٠,٧٤)-(-٠,٣٦)	٠,١٢٩٨	١٢,٩٨	١٤
٣٥٥-٣٦٥	(-٠,٣٦)-٠,٠١	٠,١٣٦٦	١٣,٦٦	١٦
٣٦٥-٣٧٥	٠,٠١-٠,٣٩	٠,١٥٥٧	١٥,٥٧	١٥
٣٧٥-٣٨٥	٠,٣٩-٠,٧٦	٠,١٢٤٧	١٢,٤٧	٨
٣٨٥-٣٩٥	٠,٧٦ - ١,١٣	٠,٠٩٤٤	٩,٤٤	١٠
٣٩٥-٤٠٥	١,١٣ - ١,٥١	٠,٠٦٣٧	٦,٣٧	٨
٤٠٥-∞	١,٥١ - ∞	٠,٠٦٦٥	٦,٥٥	٦
		١,٠٠٠٠	١٠٠,٠٠	١٠٠

في العمود الأول من هذا الجدول نضع الفئات كما هى معطاة مع تعديل واحد وهو وضع - ∞ بدلا من الحد الأدنى للفئة الأولى و + ∞ بدلا من الحد الأعلى

للفئة الأخيرة ، وذلك لأن التوزيع المعتدل هو توزيع متصل تقع قيمه بين $-\infty$ ، $+\infty$ وهذا التعديل من شأنه إدخال جميع هذه القيم دون أن يؤثر ذلك على التوزيع التكرارى المعطى .

وفي العمود الثاني نضع حدود الفئات بعد تحويلها إلى قيمها المعيارية بواسطة التعويض $\text{ع} = \frac{\text{س} - ٣٦٤,٧}{٢٦,٧}$ توطئة لاستخدام جدول المساحات أسفل المنحني المعتدل المعيارى ، فمثلا للفئة الأولى

$$\text{س} = -\infty \quad \text{تعطى} \quad \text{ع} = -\infty$$

$$\text{، س} = ٣٢٥ \quad \text{تعطى} \quad \text{ع} = \frac{٣٢٥ - ٣٦٤,٧}{٢٦,٧} = -١,٤٩$$

وهكذا بالنسبة لبقية الفئات .

وفي العمود الثالث نضع التكرارات النسبية المتوقعة $\text{ل}_\text{ر}$ التي تعبر عن احتمالات وقوع قيم المتغير ع في الفئات المناظرة ، وهذه الاحتمالات نوجدها من جدول المساحات . فمثلا للفئتين الأولى والثانية :

$$\text{ل}\,(-\infty < \text{ع} \leq -١,٤٩) = ٠,٥٠٠٠ - ٠,٤٣١٩ = ٠,٠٦٨١$$

$$\text{ل}\,(-١,٤٩ < \text{ع} \leq -١,١١) = ٠,٤٣١٩ - ٠,٣٦٦٥ = ٠,٠٦٥٤$$

وهكذا بالنسبة لبقية الفئات .

ولما كانت هذه الاحتمالات هى بمثابة التكرارات النسبية في كل فئة فإننا للمقارنة بالتوزيع المعطى نضرب كلا منها في حجم هذا التوزيع وهو هنا ١٠٠ لنحصل على التكرارات المتوقعة $\text{ق}_\text{ر}$ أى التكرارات التي نتوقعها في حالة كون المجتمع معتدلا . وهذه نضعها في العمود الرابع . أما العمود الخامس فيحمل التكرارات المشاهدة في العينة لتسهيل مقارنتها بالتكرارات المتوقعة .

ومن هذه المقارنة نشعر أن هناك مطابقة حسنة بين التوزيع التكرارى المشاهد والتوزيع التكرارى النظرى إذ لا تختلف التكرارات المشاهدة عن نظائرها المتوقعة في أغلب الفئات إلا قليلا مما يشير إلى أن المجتمع الذى أخذت منه العينة هو على الأرجح مجتمع معتدل ، وإن كان الحكم الموضوعى في ذلك يتطلب استخدام أحد الاختبارات الإحصائية المناسبة . انظر المثال (٦ – ٩) في البند (٦ – ٧ – ٢) .

(٤ – ٤) طريقة بيانية للكشف عن الاعتدالية :

هناك طريقة بيانية تستخدم كاختبار سريع للكشف عن دلالة المنحني التكرارى المشاهد ومدى انحرافه عن الاعتدالية ، ويشترط في هذه الطريقة أن تكون العينة عشوائية وكبيرة الحجم (ن أكبر من ٥٠) . وتبنى فكرة هذه الطريقة على أن التوزيع المعتدل هو توزيع متماثل ذو تفرطح معين وبالتالى فإن أهم ما ينبغى التحقق منه في توزيع تكرارى لعينة هو مدى تماثله ومدى تفرطحه بالنسبة للتوزيع المعتدل .

وكل ما تتطلبه هذه الطريقة هو تكوين توزيع التكرارات المتجمعة المئوية للتوزيع التكرارى المعلوم ثم رسم النقط التي تمثل هذا التوزيع على ورق تقسيم الاحتمالات . فإذا وقعت هذه النقط على وجه التقريب على خط مستقيم يمكن توفيقه بالعين دل ذلك على أن المجتمع هو على الأرجح مجتمع معتدل ، وإلا فهو ليس كذلك . راجع الخاصة (هـ) من البند (٤ – ١) .

ولهذه الطريقة فائدة أخرى ، وهى أنه إذا ظهر لنا أن المجتمع معتدل أو قريب من الاعتدال فإننا نستطيع تقدير الوسط الحسابي والانحراف المعيارى لهذا المجتمع من المستقيم الذى وفقناه كالآتي :

(أ) الوسط الحسابي (الوسيط في الواقع) يقدر بالإحداثي السيني للنقطة التي على الخط المستقيم التي إحداثيها الصادى ٥٠ .

(ب) الانحراف المعيارى يقدر بنصف الفرق بين الإحداثيين السينيين للنقطتين اللتين إحداثياهما الصاديان ١٥,٩ ، ٨٤,١ .

مثال (٤ – ٥) :

استخدم الطريقة البيانية لاختبار اعتدالية المجتمع الذى سحبت منه العينة المذكورة في المثال (٤ – ٤) ، وإذا رأيت أن المجتمع معتدل فأوجد تقديراً لوسطه الحسابي وانحرافه المعيارى .

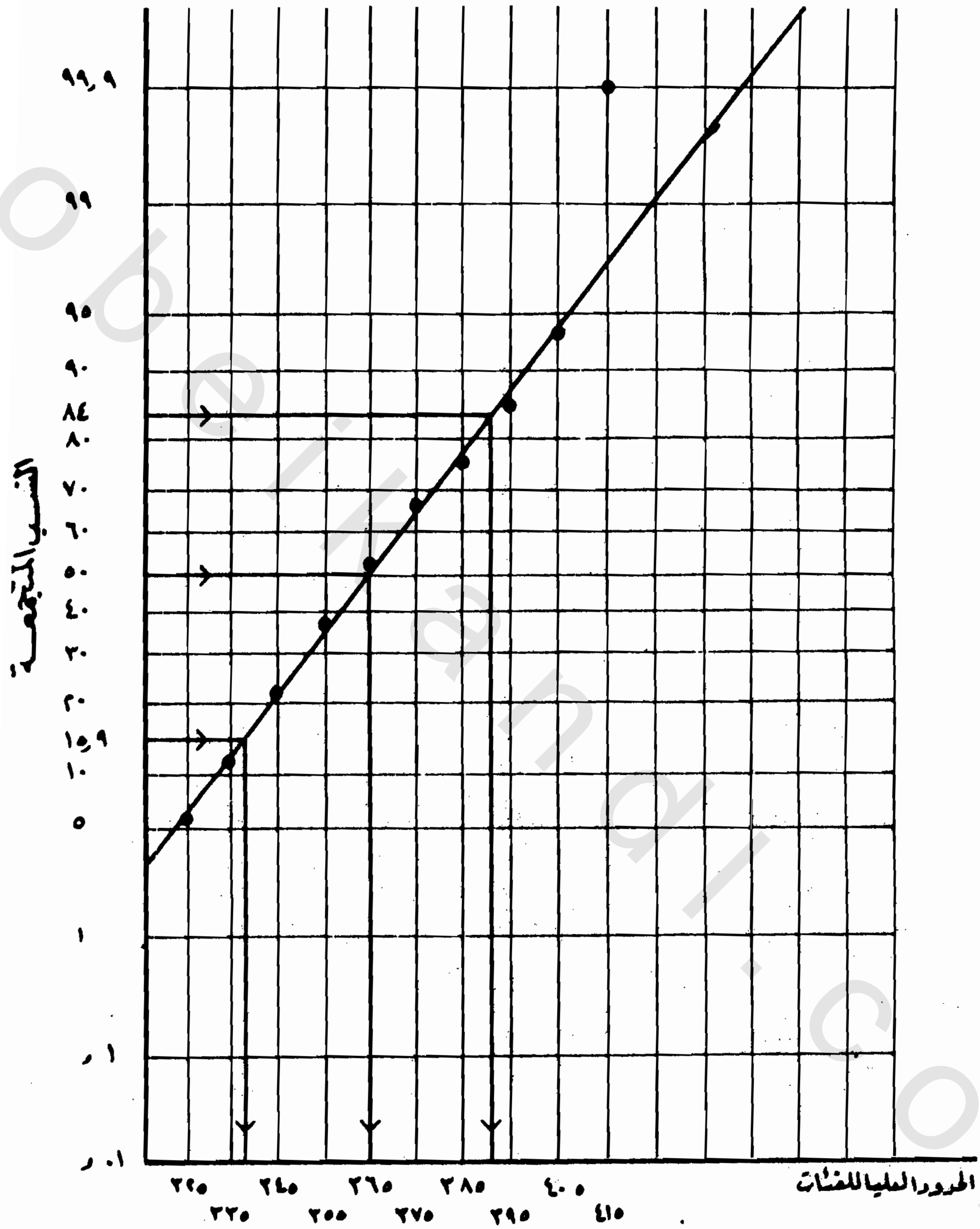
الحل :

نبدأ بإنشاء توزيع التكرارات المتجمعة المئوية كما في الجدول (٤ – ٢) الآتي :

الجدول (٤ – ٢)

الحدود العليا للفئات	التكرار المتجمع	التكرار المتجمع ٪
⩾ ٣٢٥	٦	٦
⩾ ٣٣٥	١٢	١٢
⩾ ٣٤٥	٢٣	٢٣
⩾ ٣٥٥	٣٧	٣٧
⩾ ٣٦٥	٥٣	٥٣
⩾ ٣٧٥	٦٨	٦٨
⩾ ٣٨٥	٧٦	٧٦
⩾ ٣٩٥	٨٦	٨٦
⩾ ٤٠٥	٩٤	٩٤
⩾ ٤١٥	١٠٠	١٠٠

نرسم النقط (٣٢٥ ، ٦) ، (١٢,٣٣٥) ، ... ، (٤١٥ ، ١٠٠) على ورق تقسيم الاحتمالات المعتدلة لنحصل على الشكل (٤ – ٦) الآتى :



الشكل (٤ – ٦) توزيع التكرارات النسبية المتجمعة لبيانات المثال (٤ – ٤)

مرسوماً على ورق تقسيم الاحتمالات المعدلة

بالتأمل في هذا الشكل نجد أن النقط تكاد تقع على خط مستقيم مما يشير إلى اعتدالية التوزيع . من الخط المرسوم نقدر الوسط الحسابي والانحراف المعيارى للمجتمع كالآتي :

الوسط الحسابي = ٣٦٥

الانحراف المعيارى = $\frac{١}{٢}$ (٣٩٢ - ٣٣٩) = ٢٦,٥

(٤ - ٥) معالجة عدم اعتدالية التوزيع :

في التحليل الإحصائي للبيانات يتطلب الأمر في بعض الحالات أن يكون المجتمع معتدلا ، فإذا لم يكن المجتمع معتدلا نبحث عن تحويل مناسب يجعله معتدلا أو قريباً من الاعتدال . ومن أكثر التحويلات استخداماً في هذا الصدد التحويل المسمى بالتحويل اللوغاريتمى logarithmic transformation الذى يحول المتغير س الذى لدينا إلى متغير ص حيث ص = لو س . ولا بأس من أخذ اللوغاريتمات العادية أى ذات الأساس ١٠ . وفي كثير من الحالات يفلح هذا التحويل في تعديل التوزيعات التكرارية الملتوية إلى اليمين إلى توزيعات أكثر تماثلا وبالتالى يكون قد عالج إلى حد ما عدم اعتدالية التوزيع . هذا مع ملاحظة أن عدم الاعتدالية لا تترتب عليه نتائج وخيمة إلا إذا كان التوزيع ملتوياً التواء شديداً . على أنه إذا لم يوجد تحويل يصلح لتصحيح الاعتدالية فإننا نلجأ في تحليل البيانات إلى طرق لا تتطلب شرط الاعتدالية وهذه الطرق تسمى بالطرق غير البارامترية non-parametric methods وهى طرق لا تعتمد على فرض توزيعات محددة للمجتمعات أو المتغيرات التي ندرسها . انظر الفصل الرابع عشر .

ملاحظة عن التحويلات :

(١) يستخدم التحويل اللوغاريتمى أيضاً في تحويل نموذج من النوع الضربي مثل ص = $س_١ \, س_٢ \, س_٣$ إلى نموذج من النوع الخطى $صَ = سَ_١ + سَ_٢ + سَ_٣$

الذى هو أسهل تناولا ، وذلك بوضع ص = لو ص ، س = لو س ... الخ ، وهذا ما نفعله أحياناً في موضوع تحليل الانحدار . كما يستخدم التحويل اللوغاريتمى حين يكون الوسط الحسابي μ للمجتمع مرتبطاً ارتباطاً موجباً بالتباين σ^2 فهو يحول المتغير الذى لدينا إلى متغير آخر يكون فيه هذان الدليلان مستقلين .

(٢) هناك تحويل آخر يسمى بتحويل الجذر التربيعي square-root transformation حيث نضع ص = $\sqrt{\text{س}}$. ويستخدم هذا التحويل للبيانات التي تنتج عن العد ويكون توزيعها بواسونيا حيث يكون الدليلان μ ، σ^2 غير مستقلين (إذ نعلم أن $\mu = \sigma^2$) ويفلح هذا التحويل في جعلهما مستقلين . وإذا احتوت البيانات على أصفار يفضل استخدام التحويل ص = $\sqrt{\text{س} + \frac{1}{2}}$.

(٣) من التحويلات الشهيرة أيضاً التحويل الزاوى angular transformation حيث نضع ص = $\text{حا}^{-1}\sqrt{\text{س}}$ ويستخدم حين تكون البيانات مؤلفة من نسب مئوية . وفي توزيع ذى الحدين الذى دليلاه ن ، ح نعلم أن الوسط الحسابي μ = ن ح والتباين σ^2 = ن ح (١ − ح) وبالتالى فإن التباين يكون دالة في الوسط الحسابي . إن التحويل الزاوى يوقف هذه الدالية . إلا أنه حين تكون النسب واقعة بين ٠,٣٠ ، ٠,٧٠ فإن هذا التحويل لا يكون له ضرورة .

(٤) إن التحويلات الثلاثة السابقة تغير العلاقة الدالية بين المتغيرات أى تغير النموذج الرياضى إلى نموذج آخر . غير أن هناك تحويلات لا تفعل هذا وإنما تستهدف تقنين المتغيرات عن طريق :

(أ) استبعاد وحدات القياس وذلك بالتحويل إلى مقياس نسبى لا يعتمد على وحدات القياس ،

(ب) جعل القيم الناتجة عن التحويل فى المجموعات المختلفة تتساوى فى أوساطها الحسابية وفى تبايناتها .

وأشهر هذه التحويلات يأخذ الصورة الآتية المسماه بالصورة المعيارية :

$$ص_ر = \frac{س_ر - \bar{س}}{ع}$$

حيث $\bar{س}$ متوسط قيم س ، ع انحرافها المعيارى .

ونظرا لأن هذا المقياس نسبى فإن القيم الصادية الناتجة تكون خالية من أى وحدة قياس ، كما أن هذا التحويل إذا أجرى على قيم المتغير فى أى مجموعة فإنه يحول هذه القيم إلى قيم متوسطها يساوى صفرا وتباينها يساوى الواحد الصحيح .

تمارين (٤)

(١) للتوزيع المعتدل المعيارى وباستخدام جدول المساحات

(أ) أوجد كلا من الاحتمالات الآتية :

ل (١,٥ $\geqslant$ ع > ٠) ، ل (ع $\leqslant$ ١,٦٥) ، ل (ع < − ٠,٧٤)

، ل (١,٢٥ $\geqslant$ ع > − ٢,٠٥) .

(ب) أوجد قيمتي أ ، ب بحيث ل (ع < ب) = ٠,٩٠ ، ل (ع > أ) = ٠,٤٨

(جـ) أثبت أن ل (١,٩٦ $\geqslant$ ع > − ١,٩٦) = ٠,٩٥ ، ل (ع $\leqslant$ ١,٦٤) = ٠,٠٥

، ل (٢,٥٨ $\geqslant$ ع > − ٢,٥٨) = ٠,٩٩ ، ل (ع $\leqslant$ ٢,٣٣) = ٠,٠١

(٢) للتوزيع المعتدل الذى وسطه الحسابي ٥٠ وانحرافه المعيارى ٥ أوجد كلا من ل (٦٠,٥ $\geqslant$ س > ٣٠,٥) ، ل (٥٥ $\geqslant$ س > ٤٥) .

(٣) في مجتمع معين المعروف أن نسب الذكاء تتوزع توزيعاً معتدلا وسطه الحسابي ١٠٤,٦ وانحرافه المعيارى ١٣,١٥ .

(أ) أوجد نسبة الأفراد الذين تقع نسب ذكائهم بين ٩٠ ، ١٢٠

(ب) أوجد نسبة الأفراد الذين تزيد نسب ذكائهم عن ١١٠

(٤) التوزيع الآتي هو التوزيع التكراري لعينة عشوائية مأخوذة من مجتمع ما .

٢٧,٥–	٣٢,٥ –	٣٧,٥–	٤٢,٥–	٤٧,٥–	٥٢,٥–	٥٧,٥–	٦٢,٥–
٥	٦	١٣	٨	١٦	١٤	٧	١

(أ) أثبت أن الوسط الحسابي يساوى ٤٧,٠٧١ وأن الانحراف المعيارى يساوى ٨,٩٤٨ .

(ب) وفق توزيعاً معتدلا لهذا التوزيع واذكر رأيك فيما إذا كان بالامكان اعتبار أن المجتمع معتدل .

(جـ) استخدم الطريقة البيانية لاختبار اعتدالية المجتمع .

(٤ – ٦) تقريب توزيع ذى الحدين بتوزيع معتدل :

فى البند (٣ – ٤ – ١) رأينا أنه إذا كان س متغيرا عشوائيا ذا توزيع ذى حدين دليلاه ن ، ح يجوز تقريب هذا التوزيع بتوزيع بواسون متوسطه يساوى متوسط توزيع ذى الحدين بشرط أن يكون حجم العينة ن كبيرا وأن يكون الاحتمال ح صغيرا حيث يكون التوزيع ملتويا إلى اليمين . يجوز تحت شروط أخرى تقريب توزيع ذى الحدين : حد (ن ، ح) بتوزيع معتدل متوسطه μ = ن ح أى يساوى متوسط توزيع ذى الحدين ، وتباينه σ^2 = ن ح ك أى يساوى تباين ذى الحدين بشرط أن تكون ن كبيرة وأن تكون ح قريبة من العدد $\frac{1}{2}$ حيث يكون التوزيع متماثلا بالتقريب . تحت هذين الشرطين يقترب توزيع الإحصاءة

$$ع = \frac{س - ن ح}{\sqrt{ن ح ك}}$$ حيث ك = ١ – ح

من التوزيع المعتدل المعيارى : مع (١،٠) كلما زاد العدد ن .

ومن الناحية العملية وجد أن هذا التقريب يكون جيداً أى يمكن التجاوز عن الخطأ الناشيء عنه إذا توفر أحد الشرطين الآتيين :

(أ) إذا كانت كل من ن ح و ن ك أكبر من خمسة

أو (ب) إذا كانت ن ≤ ١٠ ومعامل الالتواء أصغر من ٠,٢ .

ونظرا لأن توزيع ذى الحدين هو توزيع لمتغير وثاب بينما التوزيع المعتدل هو توزيع لمتغير متصل فإننا لعلاج ذلك فى عملية التقريب يجب أن تعتبر أن كل قيمة س من قيم المتغير ذى الحدين ممتدة نصف وحدة من اليسار ونصف وحدة من اليمين فمثلا العدد س = ٤ نعتبر أنه الفترة (٣,٥ ، ٤,٥) ، والعدد س= ٤,٧ نعتبر أنه الفترة (٤,٦٥ ، ٤,٧٥) ، وإذا أردنا مثلا ايجاد الإحتمال ل (١٠ ≤ س ≤ ٣) لتوزيع ذى حدين فإننا نحسب الاحتمال التقريبي ل(١٠,٥ ≤س≤ ٢,٥) باستخدام جداول التوزيع المعتدل .

مثال (٤ - ٦)

ألقيت حجرة نرد منتظمة عشوائيا ١٠ مرات . أوجد احتمال الحصول على الصورة في ٣ أو ٤ أو ٥ أو ٦ مرات .

الحـل :

نظرا لأن حجرة النرد منتظمة والرمية عشوائية فإن توزيع عدد مرات ظهور الصورة هو توزيع ذو حدين : حد (١٠ ، ٠,٥) ودالة كتلة الاحتمال تكون كالآتى :

د (س) = ق^١٠_س (٠,٥)^س (٠,٥)^١٠-س = (٠,٥)^١٠ ق^١٠_س

حيث س = ٠ ، ١ ، ٢ ، ... ، ١٠

∴ ل (س = ٣ أو ٤ أو ٥ أو ٦) = د (٣) + د(٤) + د(٥) + د(٦)

$$= (٠,٥)^{١٠} \left(ق^{١٠}_{٣} + ق^{١٠}_{٤} + ق^{١٠}_{٥} + ق^{١٠}_{٦} \right)$$

$$= ٠,٧٧٣٤$$

الحل التقريبى :

بما أن الشرط (ب) متوفر إذ أن ن= ١٠ ⩾١٠ ومعامل الالتواء= صفر< ٠,٢ (مع ملاحظة أن التوزيع متماثل تماما لأن ح = $\frac{١}{٢}$) يمكن التقريب بتوزيع معتدل متوسطه μ= ن ح = ١٠×٠,٥= ٥ وتباينه $\sigma^{٢}$ = ن ح ك = ١٠×٠,٥×٠,٥= ٢,٥ واذن انحرافه المعيارى يساوى $\sqrt{٢,٥}$ = ١,٥٨. ويكون للمتغير

$$ع = \frac{س - ٥}{١,٥٨}$$

توزيع معتدل معيارى على وجه التقريب . وهنا نعتبر أن العدد ٣ هو الفترة (٢,٥ ، ٣,٥) ، وأن العدد ٤ هو الفترة (٣,٥ ، ٤,٥) وهكذا ... ويكون المطلوب إيجاد الاحتمال ل(٦,٥ >س>٢,٥) فى التوزيع المعتدل مع (٥ ، ١,٥٨)

بوضع س= ٢,٥ نجد أن ع = $\frac{٢,٥ - ٥}{١,٥٨}$ = ـ ١,٥٨٢

بوضع س= ٦,٥ نجد أن ع = $\frac{٦,٥ - ٥}{١,٥٨}$ = ٠,٩٥

∴ل(٦,٥>س>٢,٥) = ل (٠,٩٥>ع> - ١,٥٨٢) = ٠,٧٧١٨

ومن الواضح أن هذه القيمة قريبة جدا من القيمة المضبوطة ٠,٧٧٣٤

(ثانيا) التوزيع المعتدل اللوغاريتمى

إن التوزيع المعتدل اللوغاريتمى هو مثال آخر لنماذج الاحتمال المتصلة ، وقد سمى كذلك لأن التحويل ص = لو س يحول المتغير الذى يصفه هذا التوزيع إلى متغير ذى توزيع معتدل . ومن الظواهر التى يصلح لها هذا النموذج بعض الظواهر الجيولوجية كتلك المتعلقة بأوزان وأعداد بعض أنواع الصخور الرسوبية ، وبعض الظواهر الاقتصادية كدخول الأفراد وخاصة الدخول ذوات القيم الصغيرة .

ويعرّف هذا التوزيع بواسطة دالة كثافة الاحتمال الآتية :

$$د\,(س) = \frac{١}{س\sqrt{٢\,ط\,\theta^{٢}}}\; هـ^{\frac{١}{٢\theta^{٢}}(لو\,س - \mu)^{٢}} \qquad س > ٠ \qquad (٤)$$

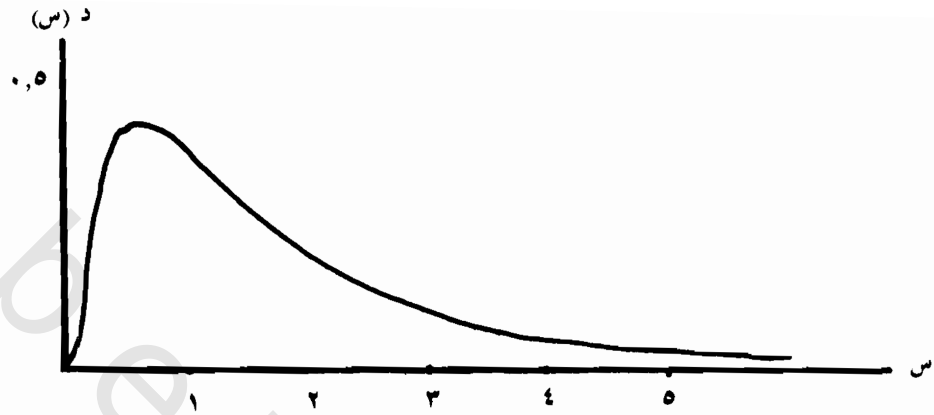
حيث اللوغاريتم للأساس هـ وحيث μ ، θ بارامتران إذا علمت قيمتاهما تحدد التوزيع تحديدا تاما .

(٤ – ٦) بعض خصائص التوزيع

(ا) المنحنى الممثل للدالة (٤) هو منحنى ذو قمة واحدة وملتوى إلى اليمين ، ويختلف شكله باختلاف قيمتى البارامترين μ ، θ ، فمثلا يأخذ الشكل المبين بالشكل (٤–٧) حين تأخذ μ القيمة صفر وتأخذ θ القيمة ١ .

(ب) التحويل ص = لو س حيث اللوغاريتم للأساس هـ يحول المتغير س إلى متغير ص له توزيع معتدل وسطه الحسابى μ وانحرافه المعيارى θ ، وبالتالى يكون المتغير

$$ع = \frac{لو\,س - \mu}{\theta}$$



الشكل (٤-٧)

المنحنى المعتدل اللوغاريتمى: ٧=٠ و θ = ١

هو متغير معتدل معيارى (وسطه الحسابى صفر وتباينه ١) . وبناء على ذلك يمكن استخدام جدول المساحات أسفل المنحنى المعتدل المعيارى ، وهو الجدول (٦) بملحق هذا الكتاب ، فى إيجاد الاحتمالات والقيم الحرجة للمتغير س كما فى المثالين الآتيين :

مثال (٤-٦)

إذا كان س متغيرا معتدلا لوغاريتميا دليلاه ٧ = ١ ، θ = ٢ فأوجد الاحتمال ل(س < ٣٥,٥) .

الحل :

بوضع ص = لو(س) يكون للمتغير ع = $\frac{\text{لو س} - ١}{٢}$ توزيع معتدل معيارى.

ل (س < ٣٥,٥) = ل (لو س < لو ٣٥,٥) = ل $\left(\frac{\text{لو س} - ١}{٢} < \frac{\text{لو ٣٥,٥} - ١}{٢} \right)$

= ل $\left(\text{ع} < \frac{٣,٥٦٩٥ - ١}{٢} \right)$ = ل (ع<١,٢٨٤٨)

= ٠,٨٩٩٧ .. (من جدول المساحات أسفل المنحنى المعتدل)

مثال (٤-٧)

إذا كان س متغيرا معتدلا لوغاريتميا دليلاه μ = ٢ ، θ= ٠,٥ . فأوجد قيمة ا بحيث ل (س < ا) = ٠,٩٠

الحـل :

$$\text{ل}\,(\text{س} < \text{ا}) = \text{ل}\,(\text{لو س} < \text{لو ا}) = \text{ل}\left(\frac{\text{لو س} - ٢}{٠,٥} < \frac{\text{لو ا} - ٢}{٠,٥}\right)$$

$$= \text{ل}\,(\text{ع} < \text{ع}) = ٠,٩٠$$

من جدول المساحات نجد أن ع = ١,٢٨ $\therefore \frac{\text{لوا} - ٢}{٠,٥}$ = ١,٢٨

$\therefore$ لوا = ٢,٦٤ $\therefore$ ا = ١٤,٠١٣