

## الفصل الرابع

### التوزيع المعتدل والتوزيع المعتدل اللوغاريتمي

### 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 المنزلي - أطوال وموازين الأطفال عند الولادة - مقادير الدسم في الزبد الناتج من بعض أنواع الأبقار ... فضلا عن أنه يلعب دورا كبيرا في بناء نظرية الاحتمال والنظرية الإحصائية .

وقد سمي هذا التوزيع بالتوزيع المعتدل ( أو المعتاد أو الطبيعي ) لأنه كان يظن فيما مضى أن أية بيانات عن ظواهر الحياة ينبغي أن تمثل 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 الألماني كارل فردريك جاوس ( ١٧٧٧ - ١٨٥٥ ) لئلا يستنتج التوزيع رياضيا كتوزيع احتمال أخطاء القياس وكان لذلك يسميه

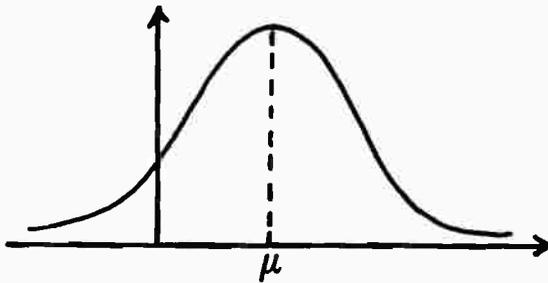
بالقانون الطبيعي للأخطاء normal law of errors ، ويسمى أيضا بتوزيع جاوس - لابلاس اعترافا بجميل العالم الفرنسي بيير سيمون لابلاس (١٧٤٩ - ١٨٢٧) الذى كان أول من اكتشفه وأثبت صلاحيته كنموذج لكثير من الظواهر . ويعرّف هذا التوزيع بواسطة دالة كثافة الاحتمال المعطاة بالقاعدة :

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

ولهذه الدالة بارامتران هما  $\mu$  ،  $\sigma$  ، تعبر  $\mu$  عن الوسط الحسابى وتعبر  $\sigma$  عن الانحراف المعياري للتوزيع . ويتحدد التوزيع تماما إذا علمت قيمتا هذين الدليلين ، ولذا نرمز له بالرمز مع  $(\sigma, \mu)$  . والمتغير العشوائى  $x$  الذى له هذه الدالة يسمى بالمتغير المعتدل .

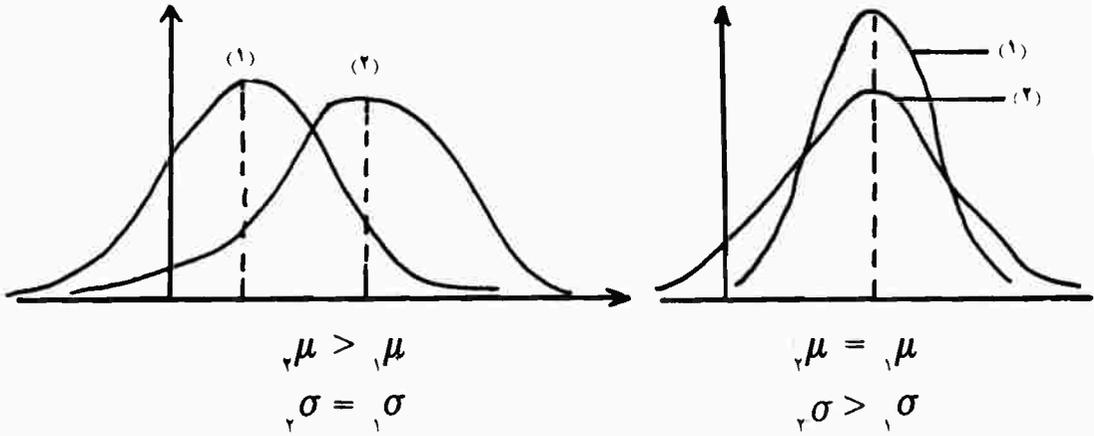
#### (٤ - ١) بعض خواص التوزيع

(أ) المنحنى الذى يمثل الدالة (١) يسمى بالمنحنى المعتدل وهو منحنى ذو قمة واحدة ومتماثل حول الخط  $x = \mu$  . ومساحة المنطقة الواقعة بين هذا المنحنى ومحور السينات تساوى الواحد الصحيح كما هو الحال لمنحنى أى توزيع احتمال . ومن التماثل نجد أن الخط  $x = \mu$  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  $\mu$ ،  $\sigma$  عدد غير منتهى من القيم ، ولذلك هناك عدد غير منتهى من التوزيعات المعتدلة . هذا مع ملاحظة أن  $\mu$  هو بارامتر موضع location parameter أما  $\sigma$  فهو بارامتر شكل shape parameter كما يتبين من الشكل (٤ - ٢) الآتي :



الشكل (٤ - ٢)

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

(د) إذا كانت  $s$  ترمز إلى قيم متغير معتدل  $s$  وسطه الحسابي  $\mu$  وانحرافه

المعياري  $\sigma$  فإن الصيغة :

$$(٢) \quad \frac{\mu - s}{\sigma} = z$$

تحول المتغير  $s$  إلى متغير  $z$  له توزيع معتدل وسطه الحسابي صفر وانحرافه المعياري الواحد الصحيح ، ويسمى هذا المتغير بالمتغير المعتدل المعياري وسنرمز له بالرمز :

مع (٠ ، ١) .

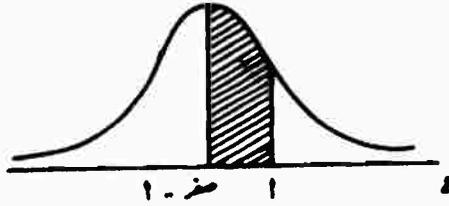
(هـ) إذا رسم توزيع التكرارات النسبية المتجمعة لمتغير معتدل على ورق الرسم

البياني ذي التقسيم الخطي ( العادي ) 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 المعياري :

نظراً لأهمية معرفة احتمالات المتغيرات المعتدلة وكثرة الحاجة إليها فقد حسبت قيم هذه الاحتمالات لمختلف فترات المتغير المعتدل المعياري  $Z$  ووضعت في جداول تعرف بجداول المساحات أسفل المنحني المعتدل المعياري ، وهذه الجداول تأخذ صوراً متعددة تؤدي إلى نفس النتائج ومنها الصورة الواردة بالجدول (٦) بملحق هذا الكتاب . وهذا الجدول يعطي المساحة تحت المنحني المعتدل المعياري وفوق الفترة بين الصفر وبعض أعداد موجبة أ متوقعة للمتغير  $Z$  ، وهذه هي المساحة المظلمة بالشكل (٤ - ٣) وهي تعبر عن الاحتمال  $L (0 < Z \leq A)$  (٣)



الشكل (٤ - ٣)

أى عن احتمال وقوع قيم المتغير  $Z$  بين العددين  $0$  ،  $A$  أى في الفترة  $(0, A)$  . ويلاحظ من تماثل المنحني أن هذا الاحتمال هو أيضاً احتمال وقوع قيم المتغير  $Z$  بين العددين  $-A$  ،  $0$  . أى :  $L (0 < Z \leq A)$  (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 ) :

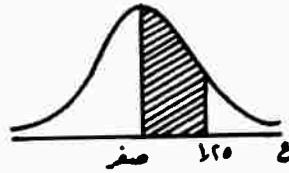
إذا علم أن توزيع درجات الطلاب في مادة ما هو توزيع معتدل وسطه الحسابي  
 $\mu = 60$  وانحرافه المعياري  $\sigma = 16$  . فأوجد باستخدام جدول المساحات  
الاحتمالات الآتية :

أولاً : ل (  $60 < س \leq 80$  )      ثانياً : ل (  $س < 80$  )  
ثالثاً : ل (  $س \geq 80$  )      رابعاً : ل (  $س \leq 76$  )  
خامساً : ل (  $75 < س \leq 85$  )

الحل :

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

$$ع = \frac{س - 60}{16}$$



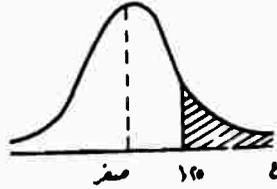
$$\text{(أولاً) بوضع } س = 60 \text{ نجد أن } ع = \frac{60 - 60}{16} = \text{صفر}$$

$$\text{وبوضع } س = 80 \text{ نجد أن } ع = \frac{60 - 80}{16} = 1,25$$

$$\text{إذن ل ( } 60 < س \leq 80 \text{ ) = ل ( } 0 < ع \leq 1,25 \text{ )}$$

$$= 0,3944$$

وذلك من الجدول مباشرة عند العدد ١,٢ الذى فى الهامش الرأسى وتحت العدد ٥ الذى فى الهامش الأفقى . وهذه النتيجة تعنى أن حوالى ٣٩٪ من الطلاب تقع درجاتهم بين ٦٠ ، ٨٠ .



(ثانياً) بوضع  $s = 80$  نجد أن  $c = 1,25$

$$\therefore (s < 80) = L(1,25 < c)$$

= مساحة المنطقة التى على يمين العدد ١,٢٥ =

$$= 0,5000 - 0,3944 =$$

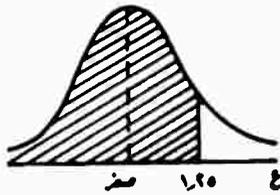
$$= 0,1056$$

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

(ثالثاً)  $L(s \geq 80) = L(c \geq 1,25)$

= مساحة المنطقة التى على يسار العدد ١,٢٥ =

$$= 0,8944 = 0,5000 + 0,3944 =$$

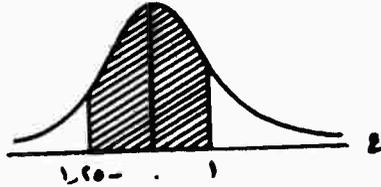


أى أن حوالى ٨٩٪ من الطلاب درجاتهم تساوى أو تقل عن ٨٠ .

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

$$( \text{رابعاً} ) \text{ بوضع } s = 40 = \text{ نجد أن } E = \frac{60 - 40}{16} = 1,25 -$$

$$\text{بوضع } s = 76 = \text{ نجد أن } E = \frac{60 - 76}{16} = 1$$



إذن ل  $(40 < s \leq 76) = L = (1 \leq E < 1,25)$   
 = مساحة المنطقة فوق الفترة  $(1, 0)$  + مساحة المنطقة فوق الفترة  $(1,25 -)$   
 $0,7357 = 0,3944 + 0,3413 = (0, 1,25 -)$

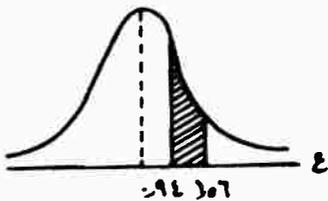
(خامساً)  $s = 75 =$  تعطي  $E = 0,9375 = 0,94$  تقريباً  
 $s = 85 =$  تعطي  $E = 1,0625 = 1,06$

إذن ل  $(75 < s \leq 85) = L = (0,94 < E \leq 1,06)$

= المساحة فوق الفترة  $(1,06, 0)$

- المساحة فوق الفترة  $(0,94, 0)$

$= 0,1142 = 0,3264 - 0,4406 =$



مثال (٤ - ٢) : مثال مشهور.

للمتغير المعتدل  $s$  الذي وسطه الحسابي  $\mu$  وانحرافه المعياري  $\sigma$  اثبت أن :

(أولاً) احتمال وقوع قيم المتغير بين العددين  $\mu \pm \sigma$  هو  $0,6826$

(ثانياً) احتمال وقوع قيم المتغير بين العددين  $\mu \pm 2\sigma$  هو  $0,9544$

(ثالثاً) احتمال وقوع قيم المتغير بين العددين  $\mu \pm 3\sigma$  هو  $0,9974$

الحل :

لكي نستخدم الجدول نحول المتغير  $s$  إلى المتغير  $z$  بواسطة التعويض :

$$\sigma / (\mu - s) = z$$

أولاً : بوضع :  $s = \mu - \sigma z$  نجد أن  $z = \frac{\mu - (\sigma - \mu)}{\sigma} = 1$

وبوضع :  $s = \mu + \sigma z$  نجد أن  $z = \frac{\mu - (\sigma + \mu)}{\sigma} = -1$

$\therefore P(\sigma - \mu < s \leq \sigma + \mu) = P(-1 < z \leq 1)$

$= P(0 < z \leq 1) \times 2$  من التماثل

$= 0,3413 \times 2 = 0,6826$

ثانياً : بالمثل نجد أن :

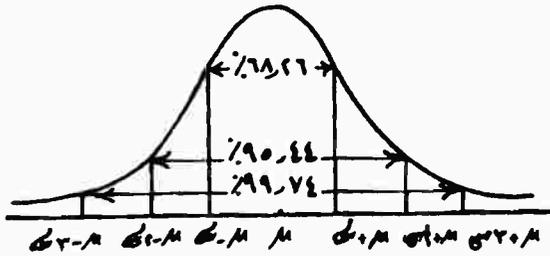
$P(\sigma \cdot 2 - \mu < s \leq \sigma \cdot 2 + \mu) = P(2 < z \leq 2)$

$= P(0 < z \leq 2) \times 2 = 0,9544$

ثالثاً :  $P(\sigma \cdot 3 - \mu < s \leq \sigma \cdot 3 + \mu) = P(3 < z \leq 3)$

$= P(0 < z \leq 3) \times 2 = 0,9974$

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



الشكل (٤ - ٤) المساحات أسفل المنحنى المعتدل

مثال (٤ - ٣) : مثال مشهور

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

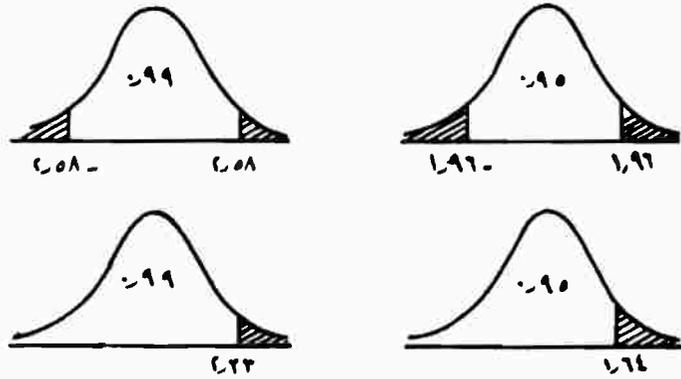
$$(١) \text{ ل } (١,٩٦ < \mathcal{E} \leq ١,٩٦) = ٠,٩٥$$

$$(٢) \text{ ل } (٢,٥٨ < \mathcal{E} \leq ٢,٥٨) = ٠,٩٩$$

$$(٣) \text{ ل } (١,٦٤ < \mathcal{E}) = ٠,٠٥$$

$$(٤) \text{ ل } (٢,٣٣ < \mathcal{E}) = ٠,٠١$$

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



الشكل (٤ - ٥) بعض القيم الحرجة للمتغير المعتدل المعارى

(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 ينبغي أن نتحقق من توفر هذا الافتراض قبل إجراء مثل هذا التحليل .

نفرض أن لدينا توزيعاً تكرارياً ذا فئات لعينة عشوائية وسطها الحسابي  $\bar{x}$  وانحرافها المعياري  $\sigma$  ونريد اختبار ما إذا كانت هذه العينة مأخوذة من مجتمع معتدل .

نتصور مجتمعاً معتدلاً له نفس الوسط الحسابي والانحراف المعياري للتوزيع التكراري المعلوم أى نأخذ  $\mu = \bar{x}$  ،  $\sigma = \sigma$  . إذا كانت العينة مأخوذة من هذا المجتمع فإن احتمال وقوع المتغير المعتدل في فئة مساوية لأى من الفئات التي ينقسم إليها التوزيع التكراري لا يجب أن يختلف كثيراً عن التكرار النسبي المشاهد في هذه الفئة . وعلى ذلك نقوم بحساب احتمالات وقوع المتغير المعتدل في جميع فئات التوزيع التكراري مستعينين في ذلك بمجدول 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 التكراري المشاهد والتوزيع التكراري النظري ، جاز لنا أن نعتبر أن المجتمع معتدل ، وإلا فهو ليس كذلك .

إن عملية إيجاد توزيع تكراري نظري بالطريقة المذكورة تسمى بعملية توفيق توزيع معتدل لتوزيع تكراري معلوم . وقد سبق أن مرت بنا فكرة التوفيق هذه في حالة كل من توزيع ذى الحدين وتوزيع بواسون . وكما سبق القول ، يعتمد الحكم الموضوعي على حسن المطابقة أو سوءها على أحد الاختبارات الإحصائية مثل اختبار  $\chi^2$  الذي سندرسه فيما بعد .

مثال ( ٤ - ٤ ) :

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

الفئة	-310	-320	-330	-340	-350	-360	-370	-380	-390	-400
التكرار	6	6	11	14	16	10	8	10	8	6

الحل :

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

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i f_i = \frac{1}{100} (\dots + 330 \times 6 + 320 \times 6 + \dots + 410 \times 6) = \frac{36470}{100} = 364,7$$

$$s^2 = \frac{1}{n-1} \left( \sum_{i=1}^k x_i^2 f_i - \frac{(\sum_{i=1}^k x_i f_i)^2}{n} \right) = \frac{1}{99} \left( \dots + 330^2 \times 6 + 320^2 \times 6 + \dots + 410^2 \times 6 - \frac{36470^2}{100} \right) = 712,89$$

$$= 712,89$$

$$s = 26,7 = \text{الانحراف المعياري}$$

نريد أن نختبر أن المجتمع الذي سحبت منه العينة هو مجتمع معتدل وسطه الحسابي 364,7 وانحرافه المعياري 26,7 . نحسب التكرارات المتوقعة كما في الجدول (4 - 1) الآتي :

الجدول (٤ - ١)

توفيق توزيع معدل للتوزيع التكرارى في المثال (٤ - ٤)

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

في العمود الأول من هذا الجدول نضع الفئات كما هي 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

$$\text{التعويض } E = \frac{364,7 - 3}{26,7} \text{ توطئة لاستخدام جدول المساحات أسفل المنحني}$$

المعتدل المعيارى ، فمثلا للفئة الأولى

$$3 - = \infty \text{ تعطى } E - = \infty$$

$$, 1,49 - = 325 = 3 \text{ تعطى } E = \frac{364,7 - 325}{26,7} = 1,49 -$$

وهكذا بالنسبة لبقية الفئات .

وفي العمود الثالث نضع التكرارات النسبية المتوقعة ل<sub>ر</sub> التي تعبر عن احتمالات وقوع قيم المتغير  $E$  في الفئات المناظرة ، وهذه الاحتمالات نوجدتها من جدول المساحات . فمثلا للفئتين الأولى والثانية :

$$L (-1,49 < E < \infty) = 0,5000 - 0,4319 = 0,0681$$

$$L (-1,11 < E < -1,49) = 0,4319 - 0,3665 = 0,0654$$

وهكذا بالنسبة لبقية الفئات .

ولما كانت هذه الاحتمالات هي بمثابة التكرارات النسبية في كل فئة فإننا للمقارنة بالتوزيع المعطى نضرب كلا منها في حجم هذا التوزيع وهو هنا ١٠٠ لنحصل على التكرارات المتوقعة ق<sub>ر</sub> 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) . وتبني فكرة هذه الطريقة على أن التوزيع المعتدل هو توزيع متماثل ذو تفرطح معين وبالتالي فإن أهم ما ينبغي التحقق منه في توزيع تكرارى لعينة هو مدى تماثله ومدى تفرطحه بالنسبة للتوزيع المعتدل .

وكل ما تتطلبه هذه الطريقة هو تكوين توزيع التكرارات المتجمعة المثوية 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 المعياري 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 المعياري يقدر بنصف الفرق بين الإحداثيين السينيين للنقطتين اللتين إحداثيها الصاديان ٩، ١٥ ، ١ ، ٨٤ .

مثال (٤ - ٥) :

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

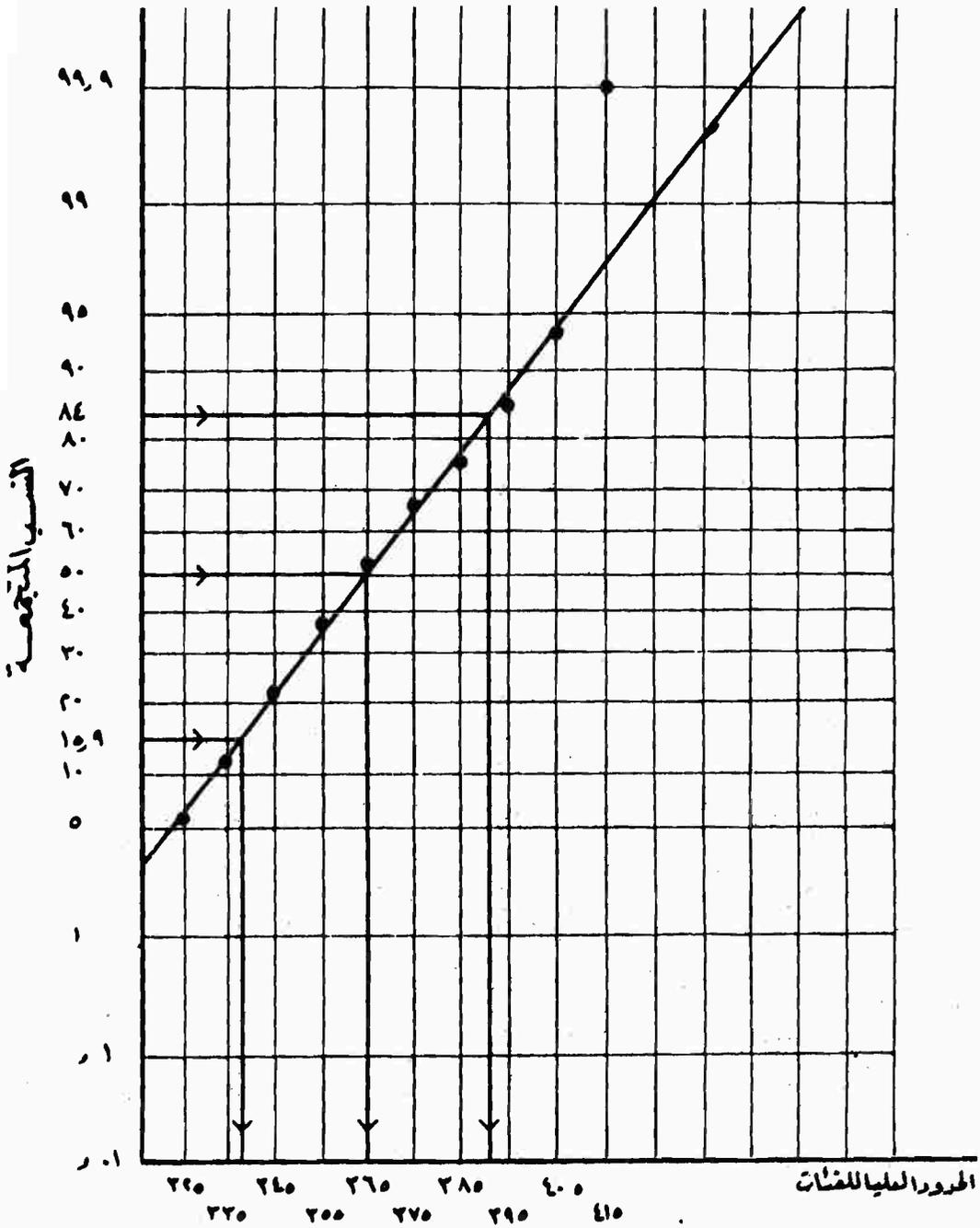
الحل :

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

الجدول (٤ - ٢)

الحدود العليا للفئات	التكرار المتجمع	التكرار المتجمع %
$325 \geq$	٦	٦
$335 \geq$	١٢	١٢
$345 \geq$	٢٣	٢٣
$355 \geq$	٣٧	٣٧
$365 \geq$	٥٣	٥٣
$375 \geq$	٦٨	٦٨
$385 \geq$	٧٦	٧٦
$395 \geq$	٨٦	٨٦
$405 \geq$	٩٤	٩٤
$415 \geq$	١٠٠	١٠٠

نرسم النقط (٦ ، ٣٢٥) ، (١٢ ، ٣٣٥) ، ... ، (٩٤ ، ٤١٥) ، (١٠٠ ، ٤١٥) على ورق تقسيم الاحتمالات المعتدلة لنحصل على الشكل (٤ - ٦) الآتي :



الشكل (٤ - ٦) توزيع التكرارات النسبية المجمعة 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 المعياري للمجتمع كالآتي :

$$\begin{aligned} \text{الوسط الحسابي} &= 365 \\ \text{الانحراف المعياري} &= \frac{1}{7} (392 - 339) = 26,5 \end{aligned}$$

#### (٤ - ٥) معالجة عدم اعتدالية التوزيع :

في التحليل الإحصائي للبيانات يتطلب الأمر في بعض الحالات أن يكون المجتمع معتدلاً ، فإذا لم يكن المجتمع معتدلاً نبحث عن تحويل مناسب يجعله معتدلاً أو قريباً من الاعتدال . ومن أكثر التحويلات استخداماً في هذا الصدد التحويل المسمى بالتحويل اللوغاريتمي logarithmic transformation الذي يحول المتغير  $v$  الذي لدينا إلى متغير  $v'$  حيث  $v' = \log v$  . ولا بأس من أخذ اللوغاريتمات العادية أي ذات الأساس ١٠ . وفي كثير من الحالات يفلح هذا التحويل في تعديل التوزيعات التكرارية المتتوية إلى البعدين إلى توزيعات أكثر تماثلاً وبالتالي يكون قد عالج إلى حد ما عدم اعتدالية التوزيع . هنا مع ملاحظة أن عدم الاعتدالية لا تترتب عليه نتائج وخيمة إلا إذا كان التوزيع متتوياً 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 وهي طرق لا تعتمد على فرض توزيعات محددة للمجتمعات أو المتغيرات التي ندرسها . انظر الفصل الرابع عشر .

#### ملاحظة عن التحويلات :

(١) يستخدم التحويل اللوغاريتمي أيضاً في تحويل نموذج من النوع الضربي مثل  $v = s_1 s_2 s_3$  إلى نموذج من النوع الخطي  $v' = s_1' + s_2' + s_3'$

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 الحسابى  $\mu$  للمجتمع مرتبطاً ارتباطاً موجباً بالتباين  $\sigma^2$  فهو يحول المتغير الذى لدينا إلى متغير آخر يكون فيه هذان الدليلان مستقلين .

(٢) هناك تحويل آخر يسمى بتحويل الجذر التربيعى square-root transformation حيث نضع  $ص = \sqrt{س}$  . ويستخدم هذا التحويل للبيانات التى تنتج عن العد ويكون توزيعها بواسونياً حيث يكون الدليلان  $\mu$  ،  $\sigma^2$  غير مستقلين ( إذ نعلم أن  $\mu = \sigma^2$  ) ويفلح هذا التحويل في جعلهما مستقلين . وإذا احتوت البيانات على أصفار يفضل استخدام التحويل  $ص = \sqrt{س + \frac{1}{4}}$  .

(٣) من التحويلات الشهيرة أيضاً التحويل الزاوى angular transformation حيث نضع  $ص = \arcsin \sqrt{س}$  ويستخدم حين تكون البيانات مؤلفة من نسب مئوية . وفي توزيع ذى الحدين الذى دليلاه  $ن$  ،  $ح$  نعلم أن الوسط الحسابى  $\mu = ن ح$  والتباين  $\sigma^2 = ن ح (ن - ح)$  وبالتالي فإن التباين يكون دالة في الوسط الحسابى . 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 في المجموعات المختلفة تتساوى في أوساطها الحسابية وفي تبايناتها .

وأشهر هذه التحويلات يأخذ الصورة الآتية المسماه بالصورة المعيارية :

$$\frac{\bar{s} - s}{\sigma} = \text{ص}$$

حيث  $\bar{s}$  متوسط قيم  $s$  ،  $\sigma$  انحرافها المعياري .

ونظراً لأن هذا المقياس نسبي 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 صفراً وتباينها يساوى الواحد الصحيح .

#### تمارين (٤)

(١) للتوزيع المعتدل المعيارى وباستخدام جدول المساحات

(أ) أوجد كلا من الاحتمالات الآتية :

$$ل (٠ < \epsilon \leq ١,٥) ، ل (١,٦٥ \leq \epsilon) ، ل (\epsilon > ٠,٧٤) ، ل (١,٢٥ \leq \epsilon < ٢,٠٥) .$$

$$(ب) أوجد قيمتي أ ، ب بحيث ل (\epsilon > ب) = ٠,٠٠,٩٠ ، ل (\epsilon < أ) = ٠,٤٨$$

$$(ج) أثبت أن ل (١,٩٦ \leq \epsilon < ١,٩٦) = ٠,٩٥ ، ل (١,٦٤ \leq \epsilon) = ٠,٠٥$$

$$ل (٢,٥٨ \leq \epsilon < ٢,٥٨) = ٠,٩٩ ، ل (٢,٣٣ \leq \epsilon) = ٠,٠١$$

(٢) للتوزيع المعتدل الذى وسطه الحسابي ٥٠ وانحرافه المعيارى ٥ أوجد كلا من ل (٦٠,٥ \leq s < ٣٠,٥) ، ل (٥٥ \leq s < ٤٥) .

(٣) فى مجتمع معين المعروف أن نسب الذكاء تتوزع توزيعاً معتدلاً وسطه الحسابي ١٠٤,٦ وانحرافه المعيارى ٠,٣,١٥

(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 التكرارى لعينة عشوائية مأخوذة من مجتمع ما .

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

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

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

(ج) استخدم الطريقة البيانية لاختبار اعتدالية المجتمع .

#### (٤ - ٦) تقريب توزيع ذى الحدين بتوزيع معتدل :

في البند (٣ - ٤ - ١) رأينا أنه إذا كان  $n$  متغيراً عشوائياً ذا توزيع ذى حدين دليلاً  $n$  ، ح يجوز تقريب هذا التوزيع بتوزيع بواسون متوسطه يساوى متوسط توزيع ذى الحدين بشرط أن يكون حجم العينة  $n$  كبيراً وأن يكون الاحتمال  $p$  صغيراً حيث يكون التوزيع ملتوياً إلى اليمين . يجوز تحت شروط أخرى تقريب توزيع ذى الحدين : حد (  $n$  ،  $p$  ) بتوزيع معتدل متوسطه  $\mu = np$  حيث أى يساوى متوسط توزيع ذى الحدين ، وتباينه  $\sigma^2 = npq$  حيث  $q = 1 - p$  حيث يكون التوزيع بشرط أن تكون  $n$  كبيرة وأن تكون  $p$  قريبة من العدد  $\frac{1}{4}$  حيث يكون التوزيع متماثلاً بالتقريب . تحت هذين الشرطين يقترب توزيع الإحصاءة

$$z = \frac{u - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

حيث  $z = u - 1$  ح

من التوزيع المعتدل المعياري : مع (١٦٠) كلما زاد العدد  $n$  .

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

- (أ) إذا كانت كل من  $h$  و  $c$  و  $n$  أكبر من خمسة  
 أو (ب) إذا كانت  $n \leq 10$  ومعامل الالتواء أصغر من ٢,٠ .

ونظراً لأن توزيع ذى الحدين هو توزيع لمتغير وثاب بينما التوزيع المعتدل هو توزيع لمتغير متصل فإننا لعلاج ذلك فى عملية التقريب يجب أن نعتبر أن كل قيمة  $s$  من قيم المتغير ذى الحدين ممتدة نصف وحدة من اليسار ونصف وحدة من اليمين فمثلاً العدد  $s = 4$  نعتبر أنه الفترة (٣,٥ ، ٤,٥) ، والعدد  $s = 4,7$  نعتبر أنه الفترة (٤,٦٥ ، ٤,٧٥) ، وإذا أردنا مثلاً إيجاد الإحتمال ل (  $10 \leq s \leq 10,5$  ) لتوزيع ذى حدين فإننا نحسب الاحتمال التقريبي ل (  $10,5 \leq s \leq 11,0$  ) باستخدام جداول التوزيع المعتدل .

#### مثال (٤ - ٦)

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

الحل :

نظراً 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 :

$$d(s) = \binom{10}{s} (0,5)^s (0,5)^{10-s}$$

حيث  $s = 0, 1, 2, \dots, 10$

$$\therefore d(s) = d(3) + d(4) + d(5) + d(6)$$

$$= (0,5) \cdot ( \binom{10}{1} + \binom{10}{2} + \binom{10}{3} + \binom{10}{4} ) = 0,7734 =$$

الحل التقريبي :

بما أن الشرط (ب) متوفر إذ أن  $n = 10 \leq 10$  ومعامل الالتواء = صفر  $< 2, 0$  ،  
 ( مع ملاحظة أن التوزيع متماثل تماما لأن  $\frac{1}{2} = \mu$  ) يمكن التقريب بتوزيع معتدل  
 متوسطه  $\mu = n \cdot p = 0,5 \times 10 = 5$  وتباينه  $\sigma^2 = n \cdot p \cdot q = 0,5 \times 10 \times 0,5 = 2,5$   
 واذن انحرافه المعياري يساوي  $\sigma = \sqrt{2,5} = 1,58$  ويكون للمتغير

$$z = \frac{s - \mu}{\sigma} = \frac{s - 5}{1,58}$$

توزيع معتدل معياري على وجه التقريب . وهنا نعتبر أن العدد 3 هو الفترة  $(0, 2)$  ،  
 $(0, 3)$  ، وأن العدد 4 هو الفترة  $(0, 3)$  ،  $(0, 4)$  وهكذا ... ويكون المطلوب إيجاد  
 الاحتمال  $P(2,5 < s < 6,5)$  في التوزيع المعتدل مع  $(0, 1,58)$

$$\text{بوضع } s = 2,5 \text{ نجد أن } z = \frac{2,5 - 5}{1,58} = -1,582$$

$$\text{بوضع } s = 6,5 \text{ نجد أن } z = \frac{6,5 - 5}{1,58} = 0,95$$

$$\therefore P(2,5 < s < 6,5) = P(-1,582 < z < 0,95) = 0,7718$$

ومن الواضح أن هذه القيمة قريبة جدا من القيمة المضبوطة 0,7734

## ( ثانيا ) التوزيع المعتدل اللوغاريتمي

إن التوزيع المعتدل اللوغاريتمي هو مثال آخر لنماذج الاحتمال المتصلة ، وقد سمي كذلك لأن التحويل  $v = \log s$  يحول المتغير الذي يصفه هذا التوزيع إلى متغير ذي توزيع معتدل . ومن الظواهر التي 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 .

ويعرف هذا التوزيع بواسطة دالة كثافة الاحتمال الآتية :

$$d(s) = \frac{1}{\theta^2} \frac{1}{s^2} \quad \text{لـ } (s > 0) \quad (4)$$

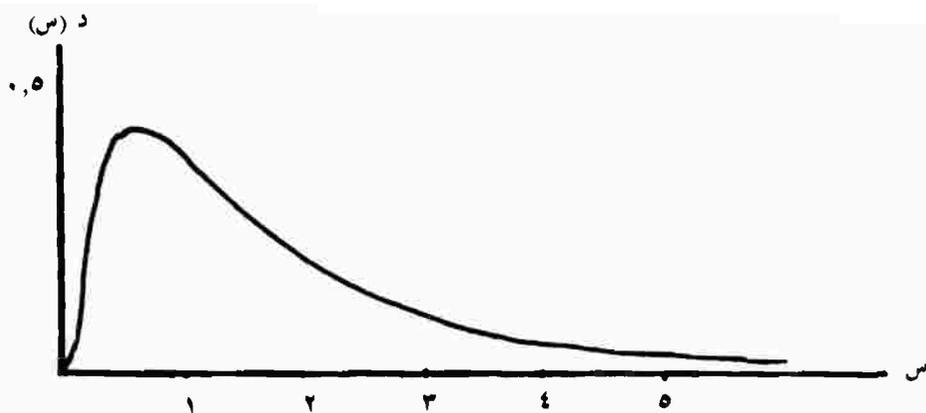
حيث اللوغاريتم للأساس  $e$  وحيث  $\theta$  ،  $\lambda$  بارامتران إذا علمت قيمتهما تحدد التوزيع تحديدا تاما .

### ( ٤ - ٦ ) بعض خصائص التوزيع

( أ ) المنحنى الممثل للدالة (٤) هو منحنى ذو قمة واحدة وملتوى إلى اليمين ، ويختلف شكله باختلاف قيمتي البارامترين  $\lambda$  ،  $\theta$  ، فمثلا يأخذ الشكل المبين بالشكل (٤-٧) حين تأخذ  $\lambda$  القيمة صفر وتأخذ  $\theta$  القيمة ١ .

( ب ) التحويل  $v = \log s$  حيث اللوغاريتم للأساس  $e$  يحول المتغير  $s$  إلى متغير  $v$  له توزيع معتدل وسطه الحسابي  $\lambda$  وانحرافه المعياري  $\theta$  ، وبالتالي يكون المتغير

$$z = \frac{\log s - \lambda}{\theta}$$



الشكل (٧-٤)

المنحنى المعتدل اللوغاريتمي:  $\gamma = 0$  و  $\theta = 1$

هو متغير معتدل معياري ( وسطه الحسابي صفر وتباينه ١) . وبناء على ذلك يمكن استخدام جدول المساحات أسفل المنحنى المعتدل المعياري ، وهو الجدول (٦) بملحق هذا الكتاب ، في إيجاد الاحتمالات والقيم الحرجة للمتغير  $s$  كما في المثالين الآتيين :

مثال (٦-٤)

إذا كان  $s$  متغيراً معتدلاً لوغاريتمياً دليلاً  $\gamma = 1$  ،  $\theta = 2$  فأوجد الاحتمال  $L(s > 35,5)$  .

الحل :

بوضع  $v = L(s)$  يكون للمتغير  $ع = \frac{1 - s^{-2}}{2}$  توزيع معتدل معياري .

$$L(s > 35,5) = L(L(s) > L(35,5)) = L\left(\frac{1 - s^{-2}}{2} > \frac{1 - 35,5^{-2}}{2}\right)$$

$$L(s > 35,5) = L\left(\frac{1 - 3,5695}{2} > ع\right) = L(ع > 1,2148)$$

$$= 0,8997 \text{ (من جدول المساحات أسفل المنحنى المعتدل)}$$

مثال (٧-٤)

إذا كان  $\theta$  متغيراً معتدلاً لوغاريتمياً دليله  $\lambda = 2$  ،  $\theta = 0,5$  ، فأوجد قيمة  $\lambda$  بحيث  $\lambda > 1$  = 0,90 .

الحل :

$$\lambda > 1 \iff (\text{لو } \lambda > \text{لو } 1) = \lambda \iff \left( \frac{\text{لو } 2 - 1}{0,5} > \frac{\text{لو } 2 - \lambda}{0,5} \right)$$

$$\lambda = (\text{ع} > \text{ع}) = 0,90$$

من جدول المساحات نجد أن  $\text{ع} = 1,28$   $\therefore \frac{\text{لو } 2 - 1}{0,5} = 1,28$

$$\therefore \text{لو } 2 = 1 \therefore 2,64 = 1 \therefore 14,013 = 1$$