الفصل الأول: العينة وتوزيعاتها الإحصائية

تمهيد:

يتم تعريف أخذ العينات على أنه عملية اختيار أعضاء معينين أو مجموعة فرعية من السكان لعمل استدلالات إحصائية منها وتقدير خصائص السكان ككل. ويستخدم أخذ العينات على نطاق واسع من قبل الباحثين في أبحاث السوق بحيث لا يحتاجون للبحث في جميع السكان لجمع رؤى قابلة للتنفيذ. وهي أيضاً طريقة ملائمة للوقت وفعالة من حيث التكلفة، وبالتالي تشكل أساس تصميم أي بحث.

على سبيل المثال، إذا أرادت شركة أدوية البحث في الآثار السلبية للدواء على سكان البلد، فمن المستحيل أن تكون قادرة على إجراء دراسة بحثية تشمل الجميع. وفي هذه الحالة، يأخذ الباحث عينة من الناس من كل تجزئة ديمغرافية ثم يجري الأبحاث عليهم مما يعطيهم ردود فعل تدل على تأثير الدواء على السكان.

1-طرق اختيار العينات

تتطلب أي دراسة بحثية في السوق نوعين أساسيين من أخذ العينات. وهم:

1. **أخذ العينة الاحتمالية:** أخذ العينة الاحتمالية هي طريقة أخذ العينات التي تختار أفراد عشوائيين من السكان من خلال وضع بعض المعايير للإختيار. تمنح معايير الاختيار هذه كل عضو فرصا متكافئة ليكون جزءًا من عينات مختلفة.
2. **اخذ العينة غير الاحتمالية:** طريقة اخذ العينة الغير احتمالية العينات تعتمد على قدرة الباحث على اختيار الأعضاء بشكل عشوائي. هذا الأسلوب في أخذ العينات ليس عملية انتقاء ثابتة أو محددة سلفا مما يجعل من الصعب علي جميع عناصر السكان الحصول على فرص متساوية لإدراجهم في عينة.

أخذ العينات الاحتمالية هي تقنية أخذ العينات التي يتم فيها اختيار عينة من مجموعة أكبر من السكان باستخدام أسلوب يقوم على نظرية الاحتمال. طريقة أخذ العينات هذه تأخذ كل فرد من السكان بعين الاعتبار وتشكل عينات على أساس عملية ثابتة. على سبيل المثال، في مجموعة سكان يبلغ عددها 1000 فرد، كل فرد من هؤلاء السكان سيحصل علي فرصه نسبتها 1000/1 ليتم اختياره ليكون جزءا من عينة. تقضي علي التحيز في مجموعة السكان وتعطي فرصة عادلة لجميع الأعضاء ليتم تضمينهم في العينة.

هناك 4 أنواع من تقنيات أخذ العينات الاحتمالية:

- **العينات العشوائية البسيطة:** واحدة من أفضل تقنيات أخذ العينات الاحتمالية التي تساعد في توفير الوقت والموارد هي طريقة "أخذ العينات العشوائية البسيطة". وطريقة جديرة بالثقة للحصول على المعلومات حيث يتم اختيار كل فرد من مجموعة السكان عشوائياً، وكل فرد يحصل علي نفس الاحتمالية ليتم اختياره ليكون جزءا من العينة

على سبيل المثال، في منظمة تضم 500 موظف، إذا قرر فريق الموارد البشرية إجراء أنشطة لبناء فريق العمل، فمن المرجح جدًا أنهم سيفضلون إجراء قرعة. في هذه الحالة، سيحصل كلّ من الـ 500 موظف علي فرصة متكافئة ليتم اختياره.

- **العينات العنقودية:** اخذ العينات العنقودية هو أسلوب يقوم فيه الباحثون بتقسيم كافة السكان إلى أقسام أو مجموعات تمثل مجموعة من السكان. يتم تحديد المجموعات وإدراجها في عينة على أساس تحديد المؤشرات الديمغرافية مثل العمر، الموقع، والجنس إلخ مما يجعل استخلاص نتائج فعالة من ردود الأفعال سهلا للغاية على صانع الدراسة الاستقصائية.

- **العينات المنتظمة:** باستخدام أسلوب أخذ العينات المنتظمة، يتم اختيار أعضاء عينة على فترات منتظمة لعدد السكان. يتطلب تحديد نقطة بداية للعينة ولحجم العينة التي يمكن تكرارها على فترات منتظمة. هذا النوع من طريقة أخذ العينات له فاصل محدد مسبقاً ، وبالتالي فإن تقنية أخذ العينات هذه هي الأقل استهلاكاً للوقت.

على سبيل المثال، يعتزم الباحث جمع عينة مكونة من 500 شخص من مجموعة سكان يبلغ عددها 5000. سيتم ترقيم كل عنصر من عناصر السكان من 1−5000 وسيتم اختيار فرد من كل 10 أفراد ليكون جزءا من العينة (مجموع السكان/"حجم العينة" = 500/5000 = 10).

- **عينات عشوائية طبقية:** أخذ العينات العشوائية الطبقية هي طريقة يمكن فيها تقسيم السكان إلى مجموعات أصغر، لا تتداخل بينما تمثل السكان بالكامل معًا. أثناء أخذ العينات، يمكن تنظيم هذه المجموعات ثم سحب عينة من كل مجموعة على حدة.

2-توزيع العينات: هو توزيع لجميع القيم الممكنة لإحصاء العينة مثل الوسط الحسابي والانحراف المعياري ونسبة ظاهرة في مجتمع للعينات المسحوبة بحجم معين من مجتمع.

معالم المجتمع: هي المقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله.

إحصاءات العينة: هي المقاييس الإحصائية التي تحسب من بيانات عينة مسحوبة من مجتمع الدراسة. وللتفرقة بين المعالم والإحصاءات يجب أن نرمز لكل منها برموز تختلف عن رموز الأخرى، على سبيل المثال يرمز للمتوسط الحســـابي للمجتمع بالرمز μ، بينما يرمز للمتوسط الحسابي للعينة بالرمز X، أيضاً للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز S *وهكذا..*

وتعتبر كل إحصاءة بمثابة تقدير أو قيمة تقديرية لمعلمة المجتمع المناظرة، فيكون المتوسط الحسابي المحسوب من بيانات العينة تقدير لمعلمة المجتمع المناظرة وهي المتوسط الحسابي المسحوب منه هذه العينة وهكذا...

إن حساب قيمة المتوسط الحسابي من بيانات العينة ليس هدفاً في حد ذاته ولكن وسيلة للتعرف على المتوسط الحسابي للمجتمع موضوع الدراسة ..وهكذا الحال بالنسبة لباقي المقاييس الإحصائية التي تحسب من العينة.

إذا أخذت جميع العينات المحتملة من مجتمعها فيتوقع أن تكون متوسطات العينات موزعة بالتساوي حول متوسط مجتمع الدراسة . بعبارة أخرى، إن متوسط متوسطات العينات يساوي متوسط مجتمعها.

وهو ذلك التوزيع التكرا ري لأحد التوابع الإحصائية المحسوب من بيانات العينات العشوائية ذات الحجم الواحد والتي يمكن سحبها من مجتمع إحصائي واحد.

نفرض أننا أخذنا عينه حجمها n من مجتمع ما، ثم حسبنا بعض المقاييس الإحصائية مثل المتوسط الحسابي، التباين،... فإن كل مقياس من هذه المقاييس يعتبر متغير عشوائي في ذاته يختلف من عينة إلى أخرى – هذا المتغير العشوائي يخضع لتوزيع معين – هذا التوزيع يسمى **بتوزيع المعاينة.**

فمثلا نقول إن توزيع المعاينة للمتوسط الحسابي وهو عبارة عن توزيع جميع المتوسطات الحسابية وكذلك فإن توزيع المعاينة للتباين هو توزيع، n للعينات المأخوذة من نفس هذا المجتمع ذو الحجم ومأخوذة من نفس المجتمع، وهكذا n ... جميع التباينات المحسوبة من عينات لها نفس الحجم إذا أخذنا عينات متكررة من مجتمع

ما وقمنا بحساب متوسط كل عينة، فإننا نجد أن معظم هذه المتوسطات تختلف عن بعضها البعض، ويسمى

نظرية (1)

إذا كان X يخضع لتوزيع وسطه μ وتباينه σ^2، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن:

$$\mu_{\bar{X}} = \mu$$

$$\sigma^2{}_{\bar{X}} = \frac{\sigma^2}{n}$$

التوزيع الاحتمالي لمتوسطات العينات" **توزيع المعاينة للوسط.**

مثال تطبيقي:

بإفتراض أن لدينا مجتمع إحصائي ما حجمه N=4، والمتغير العشوائي X هو عمر الأفراد والقيم هي 18-20-22-24 سنة. وقمنا بحساب المتوسط الحسابي والانحراف المعياري لبيانات المجتمع كما يلي:

$$u = \sum x_i / N = (18+20+22+24)/4 = 21$$

$$\sigma = 2.236$$

وقمنا بسحب جميع العينات الممكنة من مجتمع الدراسة بحجم 2 وعددها 16 عينة كما يلي:

الخيار الأول	الخيار الثاني			
	18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

وتم حساب المتوسطات الحسابية لكل عينة كما يلي:

الخيار الأول	الخيار الثاني			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

تم حساب المتوسط الحسابي والانحراف المعياري لمتوسطات العينات كما يلي:

u= 21 σ = 1.58

ومن مقارنة المتوسط الحسابي والانحراف المعيار للمجتمع مع المتوسط الحسابي والانحراف المعيار لمتوسطات العينات المسحوبة من المجتمع نستنتج أنه:

إذا كان X يخضع لتوزيع وسطه μ وتباينه σ^2، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن

المتغير X يتبع التوزيع الطبيعي بمتوسط حسابي وانحراف معياري.

نظرية (2)

إذا كانت $X_1, X_2, X_3, \dots \dots X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ وتباينه σ^2، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن $\bar{X}$ يتبع التوزيع الطبيعي بوسط μ وانحراف معياري $\sigma/\sqrt{n}$ أي أن:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

يخضع لتوزيع طبيعي معياري

مثال: أخذت عينة عشوائية حجمها 9 من الأطفال حديثي الولادة في أحد المستشفيات، فإذا علم أن وزن الطفل حديث الولادة يخضع لتوزيع طبيعي 2900 غرام وانحرافه المعياري 600 غرام.

(أ) أوجد معدل وتباين والانحراف المعياري للوسط الحسابي لأوزان الأطفال في العينة.

$$X \sim N(2900, (600)^2)$$

$$\bar{X} \sim N\left(2900, \frac{(600)^2}{n}\right)$$

(ب) أوجد احتمال أن الوسط الحسابي لأوزان الأطفال في العينة يزيد عن 3100 غرام.

$$P(\bar{X} > 3100) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{3100 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z > \frac{3100 - 2900}{600/\sqrt{9}}\right)$$

$$= P\left(Z > \frac{200}{200}\right)$$

$$= P(Z > 1)$$

$$= 1 - P(Z < 1) = 1 - 0.8413 = 0.1587$$

(ج) أوجد احتمال أن الوسط الحسابي لأوزان الأطفال في العينة يقع بين 2700 و 3200 غرام.

$$P(2700 < \bar{X} < 3200) = P\left(\frac{2700-\mu}{\sigma/\sqrt{n}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{3200-\mu}{\sigma/\sqrt{n}}\right)$$
$$= P\left(\frac{2700-2900}{600/\sqrt{9}} < Z < \frac{3200-2900}{600/\sqrt{9}}\right)$$
$$= P\left(\frac{-200}{200} < Z < \frac{300}{200}\right)$$
$$= P(-1 < Z < 1.5)$$
$$= P(Z < 1.5) + P(Z < 1) - 1$$
$$= 0.9332 + 0.8413 - 1 = 0.7745$$

مثال تطبيقي:

إذا كان توزيع الرواتب الشهرية للموظفين في إحدى الشركات التجارية يقترب من التوزيع الطبيعي بمتوسط حسابي (شيقل) $\mu = 2500$ وانحراف معياري (شيقل)$\sigma = 500$. أخذت عينة حجمها 100 موظفاً من موظفي تلك الشركة.

المطلوب:

(أ) ما هو توزيع المعاينة للوسط الحسابي $\bar{X}$ ؟

(ب) ما احتمال أن متوسط الراتب الشهري للموظفين سيكون على الأقل 2550 شيقل؟

(ج) ما هو حجم العينة التي يجب سحبها ليصبح الخطأ المعياري للوسط الحسابي مساوياً 25؟

الحل

(أ) حيث أن توزيع الرواتب الشهرية للموظفين يقترب من التوزيع الطبيعي بمتوسط (شيقل) $\mu = 2500$ وانحراف معياري (شيقل) $\sigma = 500$ فإن توزيع المعاينة للوسط الحسابي $\bar{X}$ يكون أيضاً طبيعياً بمتوسط (شيقل) $\mu = 2500$ وخطأ معياري:

$$\sigma_{\bar{X}} = \frac{500}{\sqrt{100}} = 50 \text{ (شيقل)}$$

(ب) **المطلوب:** إيجاد $P\left(\bar{X} \geq 2550\right)$، حيث أن توزيع المعاينة للوسط الحسابي $\bar{X}$ هو التوزيع الطبيعي بمتوسط (شيقل) $\mu = 2500$ وخطأ معياري $\sigma_{\bar{X}} = 50$، فإنه يمكن إيجاد $P\left(\bar{X} \geq 2550\right)$ وذلك بتحويل القيمة 2550 إلى قيمة معيارية كما يلي:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{2550 - 2500}{50} = 1$$

وبالتالي يصبح المطلوب إيجاد:

$$\begin{aligned} P\left(\bar{X} \geq 2550\right) &= P\left(Z \geq 1\right) \\ &= 1 - P\left(Z < 1\right) \\ &= 1 - 0.8413 = 0.1587 \end{aligned}$$

أي أن 15.87٪ من رواتب موظفي الشركة التجارية في العينة المُختارة من موظفي تلك الشركة سيكون على الأقل 2550 شيقل شهرياً.

(ج) يمكن تحديد حجم العينة المطلوب وذلك بمساواة الخطأ المعياري لـ $\bar{X}$ بالقيمة المطلوبة كما يلي:

$$\sigma_{\bar{X}} = \frac{500}{\sqrt{n}} = 25 \Rightarrow \sqrt{n} = \frac{500}{25} = 20 \Rightarrow n = 400$$

أي أنه يجب سحب عينة حجمها يساوي 400 موظف ليصبح الخطأ المعياري للوسط الحسابي مساوياً 25. هذه النتيجة متوقعة لأن الخطأ المعياري لـ $\bar{X}$ يساوي 50 عندما $n = 100$ ، وللحصول على نصف هذا الخطأ يجب أن نضرب حجم العينة في 4.

مثال تطبيقي:

إذا علمت أن عُمر نوع معين من المصابيح الكهربائية التي تنتجه إحدى الشركات يقترب من التوزيع الطبيعي بمعدل 1000 ساعة وانحراف معياري 200 ساعة.
أُخذت عينة عشوائية حجمها 100 مصباحاً كهربائياً. أوجد احتمال أن الوسط الحسابي لعمر المصباح في العينة المختارة:

(أ) أقل من 950 ساعة.

(ب) أكبر من 1100 ساعة.

(ج) يقع بين 970 و 1050 ساعة.

الحل

بافتراض أن X يرمز إلى عُمر المصباح الكهربائي وحيث أن $n=100$ فإن
$\mu_{\bar{X}}=1000,\ \sigma_{\bar{X}}=\frac{200}{\sqrt{100}}=20$.

(أ) احتمال أن الوسط الحسابي لعمر المصباح أقل من 950 ساعة

$$\begin{aligned}P\left(\bar{X}<950\right)&=P\left(\frac{\bar{X}-1000}{20}<\frac{950-1000}{20}\right)\\&=P\left(Z<-2.5\right)\\&=0.0062\end{aligned}$$

(ب) احتمال أن الوسط الحسابي لعمر المصباح أكبر من 1100 ساعة

$$\begin{aligned}P\left(\bar{X}>1100\right)&=P\left(\frac{\bar{X}-1000}{20}>\frac{1100-1000}{20}\right)\\&=P\left(Z>5\right)\\&=1-P\left(Z<5\right)\cong 1-1=0.0000\end{aligned}$$

مع ملاحظة أن أكبر قيمة لـ Z في جدول التوزيع تساوي 3.49 وتقابل احتمال 0.9998؛ لذلك $Z=5$ فإن قيمة الاحتمال المقابل تقريباً تساوي الواحد الصحيح، أي أن $P(Z<5)\cong 1$ ، وبالتالي فإن:

$$P\left(\bar{X}>1100\right)\cong 1-1=0.0000$$

(ج) احتمال أن الوسط الحسابي لعمر المصباح يقع بين 970 و 1050 ساعة.

$$\begin{aligned}P\left(970<\bar{X}<1050\right)&=P\left(\frac{970-1000}{20}<\frac{\bar{X}-1000}{20}<\frac{1050-1000}{20}\right)\\&=P\left(-1.5<Z<2.5\right)\\&=P\left(Z<2.5\right)-P\left(Z<-1.5\right)\\&=0.9938-0.0668\\&=0.9270\end{aligned}$$

نظرية (3): النهاية المركزية (تقارب التوزيعات)

إذا كانت $X_1, X_2, X_3, \dots \dots X_n$ عينة عشوائية من مجتمع وسطه μ وتباينه σ^2، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع فإن $\bar{X}$ يقترب من التوزيع الطبيعي بوسط μ وانحراف معياري $\sigma/\sqrt{n}$ كلما كبرت n $(n \geq 30)$ أي أن:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

يخضع لتوزيع طبيعي معياري

مثال: تخضع أوزان علب سائل غسل الصحون من نوع معين لتوزيع معدله 1000 غرام، وانحرافه المعياري 80 غرام. إذا أخذت عينة عشوائية حجمها 48 علبة.

(أ) فما المعدل والتباين والانحراف المعياري للوسط الحسابي $\bar{X}$ لأوزان العلب في العينة؟

حجم العينة كبير $(n = 48 \geq 30)$، المعدل والانحراف المعياري للمجتمع معلومة، ولذلك فشروط نظرية (3) متحققة أي أن:

$$\bar{X} \sim N\left(1000, \frac{(80)^2}{48}\right)$$

$$\mu_{\bar{X}} = \mu = 1000 \ , \ {\sigma^2}_{\bar{X}} = \frac{\sigma^2}{n} = \frac{(80)^2}{48} = \frac{6400}{48} \approx 133.33$$

$$\sigma_{\bar{X}} = \sqrt{{\sigma^2}_{\bar{X}}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{48}} \approx \sqrt{133.33} = 11.55$$

(ب) ما احتمال أن يزيد الوسط الحسابي لأوزان العلب في العينة عن 1072 غرام؟

$$P(\bar{X} > 1072) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{1072-1000}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z > \frac{1072-1000}{80/\sqrt{48}}\right)$$
$$\approx P\left(Z > \frac{72}{11.55}\right)$$
$$\approx P(Z > 6.23) = 1 - P(Z < 6.23) \approx 1 - 1 = 0$$

(ج) ما احتمال أن يقل الوسط الحسابي عن 980 غرام؟

$$P(\bar{X} < 980) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{980-\mu}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z < \frac{980-1000}{80/\sqrt{48}}\right)$$
$$\approx P(Z < -1.73) = 1 - P(Z < 1.73) = 1 - 0.9582$$
$$= 0.0418$$

مثال تطبيقي:

إذا علمت أن وزن البيض يخضع لتوزيع وسطه 50 جرام وانحرافه المعياري 10 جرام. إذا قررت وزارة التموين رفض كل طبق بيض إذا نقص وزنه عن 1.41 كيلوجرام. **المطلوب**. ما هي نسبة أطباق البيض المرفوضة مع العلم بأن عدد البيض في كل طبق 30 بيضة؟

الحل

إذا افترضنا أن أوزان البيض في طبق ما هي: $x_1, x_2, \dots, x_{30}$ فإن المطلوب $P\left(\sum_{i=1}^{30} x_i < 1410\right)$ ، مع ملاحظة أنه تم تحويل وحدة الوزن إلى جرام. بالقسمة على $n = 30$ (عدد البيض في كل طبق 30 بيضة) لكل من طريخ المتباينة في الاحتمال المطلوب، بالتالي نحصل على:

$$P\left(\frac{\sum_{i=1}^{30} x_i}{30} < \frac{1410}{30}\right) = P\left(\bar{X} < 47\right)$$

حيث أن حجم العينة $n = 30$ كبيراً بالتالي يمكن تطبيق نظرية النهاية المركزية ويكون للإحصاء $\frac{\bar{X} - 50}{10/\sqrt{30}}$ تقريباً توزيع طبيعي معياري وبذلك فإن:

$$\begin{aligned} P\left(\bar{X} < 47\right) &= P\left(\frac{\bar{X} - 50}{10/\sqrt{30}} < \frac{47 - 50}{10/\sqrt{30}}\right) \\ &= P\left(Z < -1.64\right) \\ &= 0.0505 \end{aligned}$$

وبذلك تقريباً 5.1٪ من أطباق البيض مرفوضة.

نظرية (4):

إذا كانت $X_1, X_2, X_3, \dots \dots X_n$ عينة عشوائية من مجتمع طبيعي وسطه μ معلوم وتباينه σ^2 غير معلوم، وكان $\bar{X}$ يمثل الوسط الحسابي للعينة ذات الحجم n والمسحوبة من هذا المجتمع، وكان S الانحراف المعياري لهذه العينة فإن:

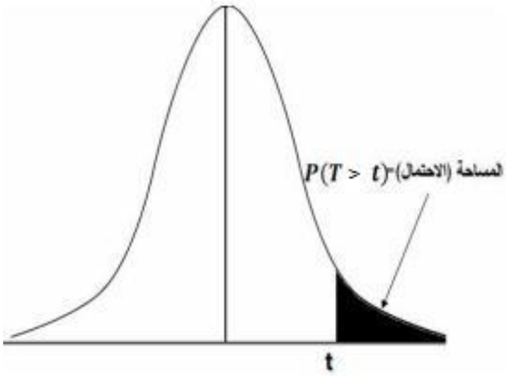
$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

يخضع لتوزيع t بدرجة حرية $n - 1$

حساب الاحتمالات تحت منحنى توزيع T

من الضروري وجود جدول يبين المساحات المختلفة تحت منحنى دالة الكثافة الاحتمالية لتوزيع T . من المعلوم أنه يوجد لكل قيمة من قيم df منحنى لدالة توزيع T ، وهذا يتطلب جدول خاص لكل قيمة من قيم df وهي عملية صعبة وتستغرق وقتاً ومجهوداً كبيرين.

لحساب الاحتمالات تحت منحنى توزيع T نحسب المساحات تحت توزيع T بعد تحديد درجات الحرية له، يوجد جداول خاصة لهذه المساحات. في جدول توزيع T يمثل العمود الأيسر درجات الحرية والخط الأفقي يمثل بعض المساحات عند قيم معينة، أما القيم الموجودة داخل الجدول فتمثل قيم T المقابلة لدرجات حرية معينة والتي تقع المساحة المطلوبة على يمين تلك القيم



منحنى دالة توزيع T

المساحة المظللة (في الشكل المقابل) تمثل $P\left[T>T(p;df)\right]$ حيث df هي درجات الحرية لتوزيع T الرمز $T[p;df]$ يمثل قيمة T التي يقع إلى يسارها مساحة مقدارها p تحت منحنى توزيع T بدرجات حرية df. جدول توزيع T يعطي قيم $T[p;df]$ لبعض قيم p الصغيرة

. مع ملاحظة أنه عندما تكون قيم p كبيرة فيتم استخدام القاعدة التالية:

$$\boldsymbol{T}[\boldsymbol{p};\ \boldsymbol{df}]=-\boldsymbol{T}[1-\boldsymbol{p};\ \boldsymbol{df}]$$

وذلك لأن منحنى توزيع T متماثل حول الصفر.

مثال تطبيقي:

بفرض أن المتغير العشوائي T يخضع لتوزيع T بدرجات حرية 20. **المطلوب** أوجد:

(أ) $P(T>1.325)$ ، أي المساحة التي تقع على يمين T=1.325

(ب) $P(T<1.72)$ ، أي المساحة التي تقع على يسار T=1.72

(ج) $P(T<-2.086)$ ، أي المساحة التي تقع على يسار.T=-2.08

(د) $P[2.086\leq T\leq 2.845]$ ، أي المساحة التي تقع بين T=2.086 ، T=2.845

(هـ) $P[-1.725\leq T\leq 1.725]$ ، أي المساحة التي تقع بين T=-1.725 ، T=1.725

(و) $T[0.005;20]$ ، أي قيمة T التي يقع على يمينها مساحة مقدارها 0.005

(ز) $T[0.90;20]$ ، أي قيمة T والتي يقع على يمينها مساحة مقدارها 0.90

الحل:

(ب) من تماثل منحنى توزيع T: $P(T<1.72)=1-P(T>1.72)$

لإيجاد $P(T>1.72)$ باستخدام جدول توزيع T نبحث أفقياً مقابل درجات الحرية 20 في العمود الأيسر عن أقرب قيمة إلى T=1.72 فنجد أن 1.725 هي أقرب قيمة لها وبالتالي فإن المساحة المقابلة لها هي 0.05، وبذلك فإن: $P(T<1.72)=1-0.05=0.95$. هذا يعني أن 95٪ من المساحة أسفل منحنى توزيع T بدرجات حرية 20 تقع على يسار القيمة 1.72

الحل

(أ) لإيجاد $P(T>1.325)$ باستخدام جدول توزيع T مباشرة نبحث أفقياً مقابل درجات الحرية 20 في العمود الأيسر عن القيمة إلى T=1.325 فنجد أن المساحة المقابلة لها هي 0.10. وبذلك فإن: $P(T>1.325)=0.10$. هذا يعني أن 10٪ من المساحة أسفل منحنى توزيع T بدرجات حرية 20 تقع على يمين القيمة 1.325

(ب) من تماثل منحنى توزيع T: $P(T<1.72)=1-P(T>1.72)$

لإيجاد $P(T>1.72)$ باستخدام جدول توزيع T نبحث أفقياً مقابل درجات الحرية 20 في العمود الأيسر عن أقرب قيمة إلى T=1.72 فنجد أن 1.725 هي أقرب قيمة لها وبالتالي فإن المساحة المقابلة لها هي 0.05، وبذلك فإن: $P(T<1.72)=1-0.05=0.95$. هذا يعني أن 95٪ من المساحة أسفل منحنى توزيع T بدرجات حرية 20 تقع على يسار القيمة 1.72

(ج) حيث أن منحنى توزيع T متماثل حول الصفر؛ بالتالي فإن $P(T<-2.086)=P(T>2.086)$ ، وبذلك: $P(T<-2.086)=0.025$.

د) **المطلوب** إيجاد $P[2.086 \leq T \leq 2.845]$:

$$P[2.086 \leq T \leq 2.845] = P[T > 2.086] - P[T > 2.845]$$
$$= 0.025 - 0.005$$
$$= 0.020$$

هـ)) **المطلوب** إيجاد $P[-1.725 \leq T \leq 1.725]$

$$P[-1.725 \leq T \leq 1.725] = P[T > -1.725] - P[T > 1.725]$$

حيث أن منحنى توزيع T متماثل حول الصفر؛ بالتالي فإن $P[T > -1.725] = P[T < 1.725]$ ، كذلك $P[T < 1.725] = 1 - P[T > 1.725]$ ، بالتالي:

$$P[-1.725 \leq T \leq 1.725] = 1 - P[T > 1.725] - P[T > 1.725]$$
$$= 1 - 2P[T > 1.725]$$
$$= 1 - 2 \times 0.05$$
$$= 0.90$$

(و) باستخدام جدول توزيع T وتحديد درجات الحرية التي تساوي 20 والمساحة 0.005 في الخط الأفقي نجد أن القيمة T=2.845 هي التي يقع علي يمينها مساحة مقدارها 0.005. أي أن $T[0.005; 20] = 2.845$

(ز) حيث أن منحنى توزيع T متماثل حول الصفر؛ بالتالي فإن $T[0.90;20]=-T[0.10;20]$. باستخدام جدول توزيع T وتحديد درجات الحرية التي تساوي 20 في العمود الأيسر والمساحة 0.10 في الخط الأفقي، نجد القيمة T=1.325 وهي القيمة التي يقع على يمينها مساحة مقدارها 0.10، وبالتالي فإن $T[0.90;20]=-T[0.10;20]=-1.325$. أي أن T= -1.325 هي القيمة التي يقع على يمينها مساحة مقدارها 0.90

مثال:

أوجد قيمة p بحيث:

(أ) $T[p;30]=2.457$ (ب) $T[p;30]=-1.310$

الحل

(أ) باستخدام جدول توزيع T وتحديد درجات الحرية التي تساوي 30 في العمود الأيسر والقيمة 2.457، نجد أن $p=0.01$.

(ب) باستخدام القاعدة $T[p;df]=-T[1-p;df]$:

$$T[p;30]=-T[1-p;30]$$

$$T[1-p;30]=1.310$$

باستخدام جدول توزيع T نجد أن $1-p=0.10 \Rightarrow p=0.90$.

مثال: إذا كانت درجات طلاب التحليل الإحصائي تأخذ شكل التوزيع الطبيعي بوسط حسابي مقداره 70 درجة. أخذت عينة حجمها 9 طلاب، ووجد أن الانحراف المعياري لعلاماتهم 11 درجات. احسب احتمال أن يزيد وسط درجاتهم عن 75 درجة.

الحل :

تباين المجتمع مجهول لذلك نستخدم توزيع t

$$
\begin{aligned}
P(\bar{X} > 75) &= P\left(\frac{\bar{X}-\mu}{S/\sqrt{n}} > \frac{75-\mu}{S/\sqrt{n}}\right) \\
&= P\left(T > \frac{75-70}{11/\sqrt{9}}\right) \qquad ,T \sim t_8 \\
&= P\left(T > \frac{5}{11/3}\right) \\
&= P\left(T > \frac{15}{11}\right) \\
&\approx P(T > 1.363) \approx P(T > 1.397) \approx 10\%
\end{aligned}
$$

إذا علمت أن توزيع أجور موظفي شركة ما يخضع للتوزيع الطبيعي بمتوسط حسابي (شيقل) $\mu = 2500$. أخذت عينة عشوائية حجمها 16 موظفاً من موظفي الشركة ووجد أن الانحراف المعياري للأجور (شيقل) $S = 200$. **المطلوب** أوجد احتمال أن يكون معدل أجورهم:

(أ) أقل من 2400 شيقل.

(ب) يزيد عن 2650 شيقل.

(ج) يتراوح بين 2350، 2600 شيقل.

الحل

بافتراض أن X يرمز إلى أجور الموظفين وحيث أن $n=16$ وتوزيع المجتمع طبيعياً وأن تباين المجتمع مجهولاً فإن المتغير العشوائي $T=\frac{\bar{X}-2500}{200/\sqrt{16}}$ لـه توزيع T بدرجات حرية $df=15$ وبذلك فإن:

(أ) احتمال أن يكون معدل أجورهم أقل من 2400 شيقل يساوي:

$$P(\bar{X}<2400)=P\left(\frac{\bar{X}-2500}{50}<\frac{2400-2500}{50}\right)$$
$$=P(T<-2)=P(T>2)$$

باستخدام جدول توزيع T وتحديد درجات الحرية التي تساوي $df=16-1=15$ في العمود الأيسر، والقيمة 2، نجد أن أقرب قيمتين لـ 2 هما 1.753، 2.131 وتقابلان الاحتمالين 0.05، 0.025 على الترتيب، بالتالي فإن

$$0.025<P(T>2)<0.05$$

القيمة الفعلية لـ $P(T>2)$ بدرجات حرية 15 تساوي 0.031973

(ب) احتمال أن يكون معدل أجورهم يزيد عن 2650 شيقل يساوي:

$$P(\bar{X}>2650)=P\left(\frac{\bar{X}-2500}{50}>\frac{2650-2500}{50}\right)$$
$$=P(T>3)$$

باستخدام جدول توزيع T وتحديد درجات الحرية التي تساوي $df=16-1=15$ في العمود الأيسر، والقيمة 3، نجد أن أقرب قيمتين لـ 3 هما 2.947، 3.286 وتقابلان الاحتمالين 0.005، 0.0025 على الترتيب، بالتالي فإن

$$0.0025<P(T>3)<0.005$$

القيمة الفعلية لـ $P(T>3)$ بدرجات حرية 15 تساوي 0.004486

(ج) احتمال أن يكون معدل أجورهم يتراوح بين 2350، 2600 شيقل يساوي:

$$P(2350 < \bar{X} < 2600) = P\left(\frac{2350-2500}{50} < \frac{\bar{X}-2500}{50} < \frac{2600-2500}{50}\right)$$

$$= P(-3 < T < 2)$$

$$= P(T > -3) - P(T > 2)$$

$$P(T > -3) = 1 - P(T < -3)$$

$$= 1 - P(T > 3)$$

$$\cong 1 - 0.005 = 0.995$$

$$P(T > 2) \cong 0.025$$

بالتالي فإن: $P(2350 < \bar{X} < 2600) \cong 0.995 - 0.025 = 0.970$

القيمة الفعلية لـ $P(-3 < T < 2)$ بدرجات حرية 15 تساوي 0.963541