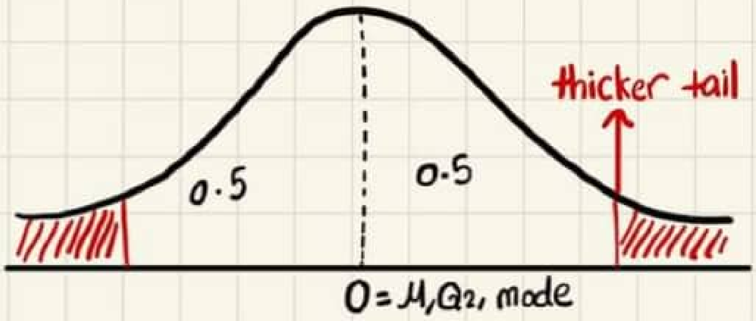
Statistics, lecture 19: اللقاء المباشر 13 + 13 فترة 24

يوتيوب عك الترتيب

$X_1, X_2, X_3, X_4 \dots X_n \sim n(\mu, \sigma^2)$ σ^2 unknown

* t-distribution:-



* كثير بيشبه الـ Normal-distribution بس الفرق انه
It has thicker tails

* كمان بيعتمد عك اشي اسمه Degrees of freedom

تعالوا نفهمها 😊

لو حكيتلك اختار ٣ ارقام بايث يطلع الوسط الحسابي 5

| ¿ | | ¿ | | ¿ |

| 2 | | 3 | | ¿ |

لنفترض انك اخترت 2,3 ، الرقم الثالث مش راح تكون حر باختيار لانك رح تعمل التالي

$$\frac{2+3+X}{3} = 5$$

$X = 10$

فبنحكي انك كنت حر باختيار رقمين وهدول الرقمين اسمهم degrees of freedom

طيب كيف نحسب degrees of freedom ؟

بشكل عام ...

$d.f = n -$ no. of Parameters

لكن في هذا الدرس ، بنتعامل فقط مع μ لذلك يصبح القانون

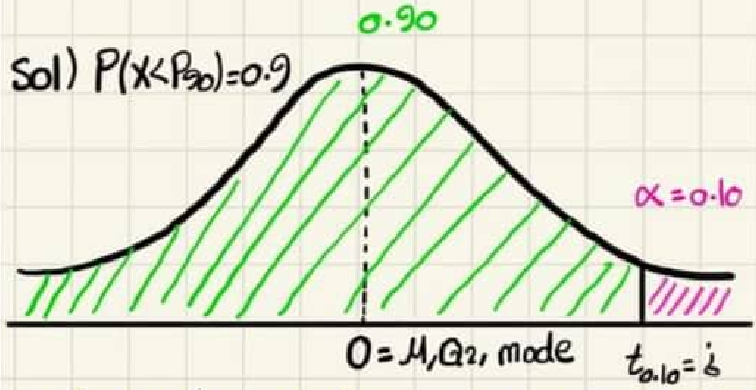
$d.f = n - 1$

وكل ما زاد حجم العينة وبالتالي d.f بصير التوزيع اقرب للـ normal وبالتالي يمكن التقريب للـ normal

* return to your book, P 332 and read the definiton 😊

e.g) If $X \sim t(7)$, then find the 90^{th} percentile ؟

Sol) $P(X < P_{90}) = 0.9$



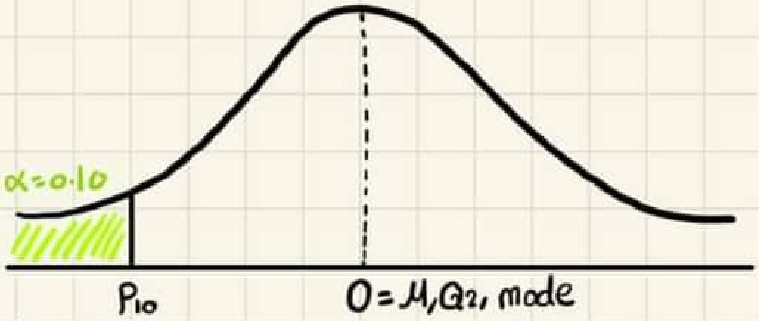
اول اشي لازم نعرف انه جدول t بتعطيني المساحة عاليمين بالتالي رح اروح عالجدول آخذ التقاطع بين $d.f = 7$ و $\alpha = 0.10$ ، ليش 0.10 ؟ لانه الرقم الي ع يساره 0.9 موجود ع يمينه 0.10

$t_{0.10} = 1.415$ "one tail, α"

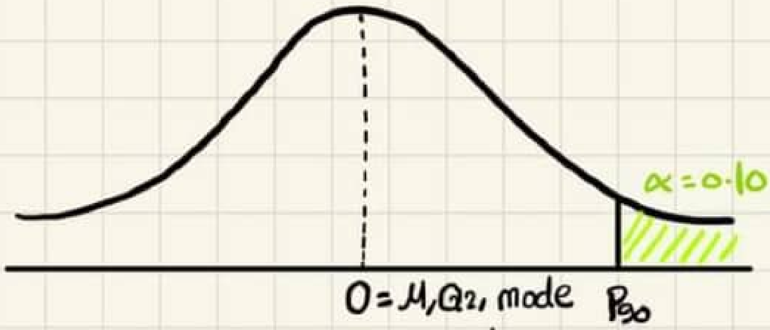
↳ t_α, t_c "critical value"

e.g) $X \sim t(7)$, find P_{10} ؟

Sol) $P(X < P_{10}) = 0.10$



* جداول ال t بتعطي المساحة باليمين وكمان ما فيها سوالب ، طيب كيف نالك في بدنا نستفيد من symmetry



لحسة طلعنا قيمتها = 1.415

بنعكس اشارة فالجواب -1.415

Note:- If $X_1, X_2, X_{\cdots n} \sim n(\mu, \sigma^2)$ then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$ or

$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim n(0, 1)$ ... If σ is known

If σ is unknown, we estimate it by S, thus

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

If σ is unknown, $n \geq 3$

$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim n(0, 1)$

يعني توزيعها t بـ

لانه n كبيرة قدرنا نقرب ل normal

e.g) $X_1, \cdots X_n \sim n(3, \sigma^2)$, If $S=2$, find the 90^{th} percentile of $\bar{X}$ ؟

Sol) $P(\bar{X} < P_{90}) = 0.90$

$P(t < \frac{P_{90} - 3}{2/\sqrt{3}}) = 0.90$

$\frac{P_{90} - 3}{2/\sqrt{10}} = 1.383$

$P_{90} = 3.87$

تأكيد كمان مرة :- القيم الي بنستخدمها من الجدول ، الكتاب بسميها critical values

((t_α , t_c))

e.g) If $X_1, X_2, X_3, \cdots X_{51} \sim n(5, \sigma^2)$ $S=2$, find the 90^{th} percentile of $\bar{X}$.

Sol) $P(\bar{X} \leq P_{90}) = 0.90$

$P(t \leq \frac{P_{90} - 5}{2/\sqrt{51}}) = 0.90$

$\frac{P_{90} - 5}{2/\sqrt{51}} = 1.299$

$P_{90} = 5.36$

كمان طريقة للحل، تقرّب لل Normal

$P(\bar{X} \leq P_{90}) = 0.90$

$P\left(Z \leq \frac{P_{90}-5}{2/\sqrt{51}}\right) = 0.90$

$\frac{P_{90}-5}{2/\sqrt{51}} = 1.28 \Rightarrow P_{90} = 5.36$

ركّز

احنا حكينا لمّا تكون $n \geq 30$ ، بنحل باستخدام التقريب لل Normal ، لكنه هون كيف زبط الحل عالطريقتين؟ زبط لانه قيمة d.f كانت 50 وهي موجودة بجداول t بس افرض طلعت معك 43 ، مش راح تلاقي قيمها بالجداول بالتالي رح تحل عالتقريب لل Normal.

* To sum up:-

If $X_1, \ldots, X_n \sim n(\mu, \sigma^2)$ then

1] If σ is known, then $\bar{X} \sim n(\mu, \frac{\sigma^2}{n})$ or $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim n(0,1)$

2] If σ is unknown then $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

for $n \geq 30$, $Z = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim n(0,1)$

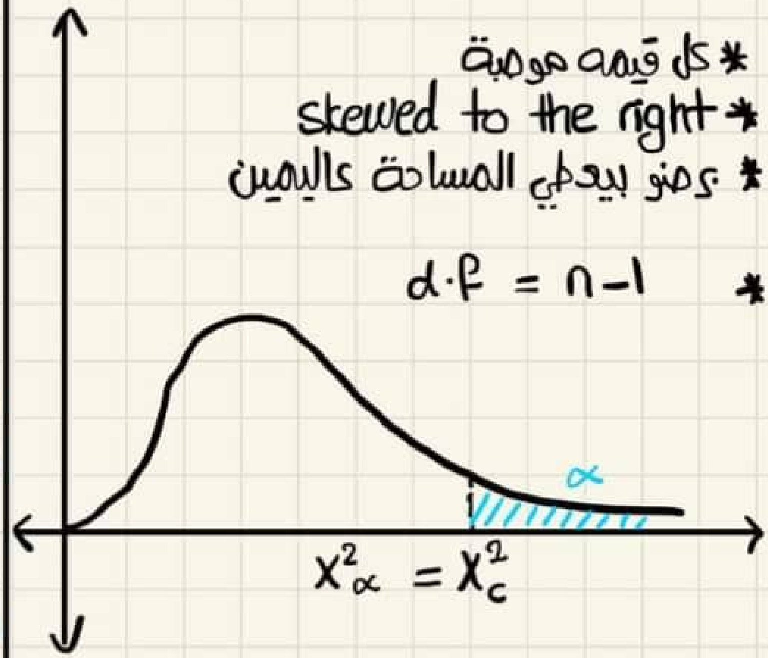
3] $\hat{P} \sim n(P, \frac{Pq}{n})$

$Z = \frac{\hat{P} - Pq}{\sqrt{Pq/n}} \sim n(0,1)$

⑤

* what is the distribution of the sample variance S^2 ؟

* Chi-squared distribution



* كل قيمه موجبة
* skewed to the right
* يعني بيعطي المساحة عاليمين
* $d.f = n-1$

* Return to your book, P 352 and read the definition.

* If $X_1, \ldots, X_n \sim n(\mu, \sigma^2)$, then

$X^2 = \frac{(n-1)}{\sigma^2} \times S^2 \sim X^2(n-1)$

e.g) If $X_1, \ldots, X_6 \sim n(\mu, 10)$, find $P(S^2 > 18.4727)$

Sol) $n = 6$, $\sigma^2 = 10$

$P(X^2 > 18.4727 \times \frac{5}{10})$

$P(X^2 > 9.2362) = 0.10$

⑥

e.g) Let $X_1, \dots X_{10} \sim n(\mu, 25)$, find the 90^{th} percentile of S^2

sol) $P(S^2 < P_{90}) = 0.90$

$P(X^2 < P_{90} \times \frac{9}{25}) = 0.90$

$P_{90} \times \frac{9}{25} = 14.684 \Rightarrow P_{90} \simeq 40.788$

****

e.g) Suppose that the weight of orange boxes are normally distributed with mean 10 kgs and standard deviation 1.5 kgs. If a no. of boxes will be loaded in a car with threshold 1000 kgs. Find the no. of boxes that will be loaded so that their total weight doesn't exceed the threshold of the car with Prob about 0.95

sol) $X_1, X_2, X_3 \dots X_n \sim n(10, 1.5^2)$

$\bar{X} \sim n(10, 1.5^2/n)$

$P(\sum_{i=1}^{n} x_i \leq 1000) = 0.95$

$P(\bar{X} \leq \frac{1000}{n}) = 0.95$

$P\left(Z \leq \frac{\frac{1000}{n} - 10}{1.5/\sqrt{n}}\right) = 0.95$

$\frac{\frac{1000}{n} - 10}{1.5/\sqrt{n}} = 1.64$

$(\frac{1000}{n} - 10) \times \frac{\sqrt{n}}{1.5} = 1.64$

let $x = \sqrt{n}$ $\therefore x^2 = n$

$(\frac{1000}{x^2} - 10) \times \frac{x}{1.5} = 1.64$

⋮

etc

$n = 97.5 \Rightarrow n \simeq 98$

e.g) Suppose that the weights of new born babies are normally distributed with mean 3kgs, a random sample of size 10 is taken and showed that it's standard deviation is 2

a) find the Prob that the sample average is below 4.16 kgs.

b) What is the 90^{th} percentile of the distribution of $\bar{X}$?

sol) $X_1, X_2, \dots X_{10} \sim n(3, \sigma^2)$

$n = 10$ $S = 2$

a) $P(\bar{X} < 4.16) = P(t < \frac{4.16 - 3}{2/\sqrt{10}})$

$P(t < 1.834) = 0.95$

b) $P(\bar{X} < P_{90}) = 0.90$

$P(t < \frac{P_{90} - 3}{2/\sqrt{10}}) = 0.90$

$\frac{P_{90} - 3}{2/\sqrt{10}} = 1.383 \Rightarrow P_{90} = 3.87$