Recall: the pivotal quantity

$$Z = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim G(0, 1)$$

can be used to form confidence intervals for μ.

Note that σ is generally unknown — it is the standard deviation of the estimator of μ. It is NOT a random variable.

Suppose we replace σ with S, the sample standard deviation (this is a random variable)

We calculate S as the square root of the sample variance

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

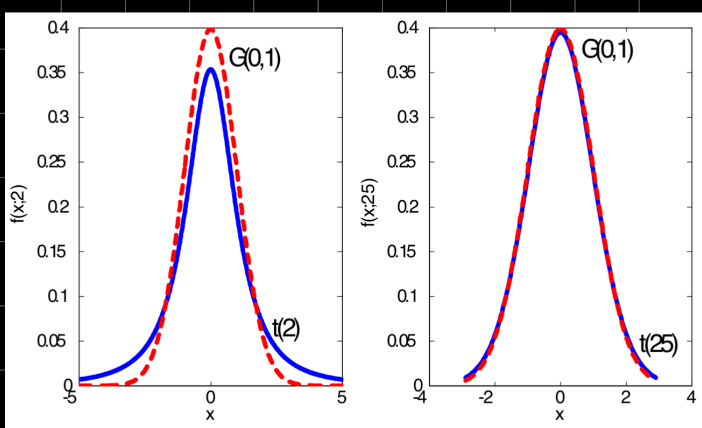
Then:

$$T = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t(\underbrace{n-1}_{\text{degrees of freedom}})$$

t-distribution

Much like the PDF of a Gaussian distribution, the PDF of the t-distribution is some fucked up equation that’s too complicated for you to ever use in this course.

The main thing that you have to remember is that T looks closer to the Gaussian for larger degrees of freedom.



Theorem 32 *Suppose $Z \sim G(0,1)$ and $U \sim \chi^2(k)$ independently. Let*

$$T = \frac{Z}{\sqrt{U/k}}$$

Then T has a ***Student's t distribution with k degrees of freedom.***

Derivation using this:

$$T = \frac{Z}{(U/k)^{1/2}} = \frac{\bar{Y} - \mu}{(U/k)^{1/2}(\sigma/\sqrt{n})}$$

Theorem

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

: pivotal quantity for estimating S^2 or S

Using $U = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
and setting $k = n-1$ degrees of freedom

$$U/k = \frac{(n-1)S^2}{\sigma^2} \cdot \frac{1}{n-1} = \frac{S^2}{\sigma^2}$$

$$\Rightarrow T = \frac{\bar{Y} - \mu}{\left(\frac{S^2}{\sigma^2}\right)^{1/2} \cdot \frac{\sigma}{n}} = \frac{\bar{Y} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

<u>Confidence intervals for the mean</u>

We want $p = P(-a \leq T \leq a)$

$$\Rightarrow P\left(-a \leq \frac{\bar{Y} - \mu}{S/\sqrt{n}} \leq a\right)$$

$$= P(-aS/\sqrt{n} \leq \bar{Y} - \mu \leq aS/\sqrt{n})$$

$$= P(-aS/\sqrt{n} - \bar{Y} \leq \mu \leq aS/\sqrt{n} - \bar{Y})$$

$$= P(\bar{Y} - aS/\sqrt{n} \leq \mu \leq \bar{Y} + aS/\sqrt{n})$$

switch inequalities when multiplying by -1

$$= P(\mu \in [\bar{Y} \pm aS/\sqrt{n}])$$

This is a 100p% confidence interval for μ.

Confidence intervals for the variance σ^2

To form a 100p% confidence interval for the variance, we want

$$p = P(a \leq U \leq b)$$

Where U is the pivotal quantity for the variance from earlier.

Note that U is bounded between a and b, not [-a, a] like in the Gaussian because U follows the chi-squared distribution, which is not symmetric.

$$\Rightarrow P\left(a \leq \frac{(n-1)S^2}{\sigma^2} \leq b\right)$$

→ sample variance

↳ variance → unknown!

a and b taken from χ^2 table

$$\Rightarrow P(a \cdot (n-1)^{-1} \cdot S^{-2} \leq \sigma^{-2} \leq b \cdot (n-1)^{-1} \cdot S^{-2})$$

$$= P\left[\frac{(n-1)S^2}{b} \leq \sigma^2 \leq \frac{(n-1)S^2}{a}\right]$$

flip inequality ⇒ switch a and b

As such, a 100p% CI for σ^2 is

$$\left[\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right]$$ ☆

Prediction Intervals for Future Observations

PQ: $$\frac{Y - \overline{Y}}{S\sqrt{1 + \frac{1}{n}}} \sim t(n-1)$$

100p% PI: $$\left[\bar{y} - as\sqrt{1+\frac{1}{n}},\ \bar{y} + as\sqrt{1+\frac{1}{n}}\right]$$

This is wider than a 100p% confidence interval

- This is centered around Y, a random variable, which has its own variability.
- A confidence interval is set around μ, which is not