7 Oct 2025 - lec 3

- EXACT DISTRIBUTION of $\hat{\beta}_1$ and $\hat{\beta}_2$

> Preliminary result
> Given $Y_1,\dots,Y_n$ independent with distribution $Y_i \sim N(\mu_i, \sigma^2)$ $i=1,\dots,n$
> and a sequence of known constants a_i, $i=1,\dots,n$,
> $\sum_{i=1}^n a_i Y_i \sim N\left(\sum_{i=1}^n a_i \mu_i,\ \sigma^2 \sum_{i=1}^n a_i^2\right)$

We have seen that $\hat{\beta}_1$ and $\hat{\beta}_2$ are linear combinations of $Y_1,\dots,Y_n$ of the form
$\hat{\beta}_1 = \sum_{i=1}^n v_i Y_i \qquad \hat{\beta}_2 = \sum_{i=1}^n w_i Y_i$
hence $\hat{\beta}_1$ and $\hat{\beta}_2$ are exactly Gaussian-distributed r.v. (see res. 1)

Moreover, the expression of the two estimators are the same we obtained with OLS.
In fact, the Gaussian linear model is a special case. Hence the properties we computed still hold.
In particular, we computed

$\mathbb{E}[\hat{\beta}_1] = \beta_1 \qquad var(\hat{\beta}_1) = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)$

$\mathbb{E}[\hat{\beta}_2] = \beta_2 \qquad var(\hat{\beta}_2) = \frac{\sigma^2}{\sum_{i=1}^n (x_i-\bar{x})^2}$

The exact distributions are then easily obtained as

$$\hat{\beta}_1 \sim N\left(\beta_1;\ \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)\right)$$
$$\hat{\beta}_2 \sim N\left(\beta_2;\ \frac{\sigma^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)$$

- EXACT DISTRIBUTION of $\hat{\Sigma}^2$

$\hat{\Sigma}^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)^2$

it is possible to show that

$\frac{n\hat{\Sigma}^2}{\sigma^2} \sim \chi^2_{n-2}$ Chi-squared with n-2 degrees of freedom

In general, for a χ^2_ν r.v., the expected value is ν

$\mathbb{E}\left[\frac{n\hat{\Sigma}^2}{\sigma^2}\right] = (n-2) \Rightarrow \mathbb{E}[\hat{\Sigma}^2] = \frac{(n-2)}{n}\sigma^2$ biased

hence again we obtain an unbiased estimator as

$S^2 = \frac{n}{n-2}\hat{\Sigma}^2 \qquad \mathbb{E}[S^2] = \frac{n}{n-2}\mathbb{E}[\hat{\Sigma}^2] = \frac{n}{n-2}\cdot\frac{n-2}{n}\sigma^2 = \sigma^2$

and $\frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{n-2}$.

Moreover, it is possible to show that $\hat{\Sigma}^2 \perp\!\!\!\perp (\hat{\beta}_1, \hat{\beta}_2)$
(hence, also $S^2 \perp\!\!\!\perp (\hat{\beta}_1, \hat{\beta}_2)$)

INFERENCE ABOUT β

We have derived the exact distributions of the estimators.
With these distributions we can test statistical hypotheses, compute confidence intervals.

Examples

Test: $\begin{cases} H_0: \beta_j = b \\ H_1: \beta_j \neq b \end{cases} \qquad \begin{cases} H_0: \beta_j = 0 \\ H_1: \beta_j > 0 \end{cases} \qquad j = 1,2$

Confidence interval, of level $1-\alpha$: $\hat{C}_j$ such that $\mathbb{P}(\hat{C}_j \ni \beta_j) = 1-\alpha \quad \forall \beta_j \in \mathbb{R}$

Recall that: $\hat{\beta}_1 \sim N(\beta_1, V(\hat{\beta}_1))$ where $V(\hat{\beta}_1) = \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)$
$\hat{\beta}_2 \sim N(\beta_2, V(\hat{\beta}_2))$ $\qquad V(\hat{\beta}_2) = \frac{\sigma^2}{\sum_{i=1}^n (x_i-\bar{x})^2}$
$\frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{n-2}$

- CONFIDENCE INTERVAL for β_j

> Preliminary result:
> If $Z \sim N(0,1)$ and $W \sim \chi^2_\nu$ independent, then $\frac{Z}{\sqrt{W/\nu}} \sim t_\nu$.
> (Student's t with ν degrees of freedom)
> $\hookrightarrow$ · symmetric distrib.
> · heavier tails than a normal
> · for large ν it is very close to a normal

First of all, we need to find a pivotal quantity.

PIVOTAL QUANTITY: a transformation of the data (and of the parameter) whose distribution does not depend on the parameter (hence is completely known).

Since $\hat{\beta}_j \sim N(\beta_j, V(\hat{\beta}_j))$, the simplest (and most intuitive) transformation is

$\frac{\hat{\beta}_j - \beta_j}{\sqrt{V(\hat{\beta}_j)}} \sim N(0,1)$

however, $V(\hat{\beta}_j)$ includes σ^2 which is unknown

In place of $V(\hat{\beta}_j)$ we use an estimate, $\hat{V}(\hat{\beta}_j) = \frac{S^2}{\sigma^2}V(\hat{\beta}_j) \qquad \left(\text{e.g. } \hat{V}(\hat{\beta}_2) = \frac{S^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)$

$T_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{V}(\hat{\beta}_j)}}$

hence random!

what is its distribution? Notice that $\hat{V}(\hat{\beta}_j)$ includes S^2 (transformation of Y)

$$T_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{V}(\hat{\beta}_j)}} = \frac{\frac{\hat{\beta}_j - \beta_j}{\sqrt{V(\hat{\beta}_j)}}}{\sqrt{\frac{\hat{V}(\hat{\beta}_j)}{V(\hat{\beta}_j)}}} \longrightarrow \hat{V}(\hat{\beta}_j) = \frac{S^2}{\sigma^2}V(\hat{\beta}_j) \Rightarrow \frac{\hat{V}(\hat{\beta}_j)}{V(\hat{\beta}_j)} = \frac{S^2}{\sigma^2}\cancel{V(\hat{\beta}_j)}\cdot\frac{1}{\cancel{V(\hat{\beta}_j)}}$$

$$= \frac{\frac{\hat{\beta}_j - \beta_j}{\sqrt{V(\hat{\beta}_j)}}}{\sqrt{\frac{S^2}{\sigma^2}\cdot\frac{(n-2)}{(n-2)}}}$$

$$= \frac{\overbrace{\frac{\hat{\beta}_j - \beta_j}{\sqrt{V(\hat{\beta}_j)}}}^{\sim N(0,1)}}{\sqrt{\underbrace{\frac{(n-2)S^2}{\sigma^2}}_{\sim \chi^2_{n-2}}\cdot\frac{1}{(n-2)}}} = \sqrt{\frac{\chi^2_{n-2}}{n-2}}$$

moreover, $\hat{\beta}_j \perp\!\!\!\perp S^2$

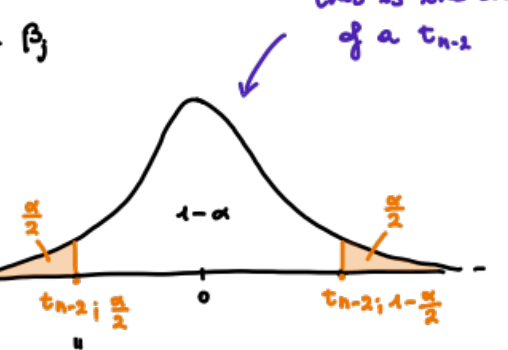
$\Rightarrow T_j \sim t_{n-2}$

CONFIDENCE INTERVAL: we want to find an interval (u,v) such that $\mathbb{P}(u < T_j < v) = 1-\alpha$ and then "isolate" the parameter to find an interval for β_j

$T_j \sim t_{n-2}$ hence

$\mathbb{P}\left(t_{n-2;\frac{\alpha}{2}} < T_j < \underbrace{t_{n-2;1-\frac{\alpha}{2}}}_{\text{quantile } 1-\frac{\alpha}{2} \text{ of a } t_{n-2} \text{ distrib.}}\right) = 1-\alpha$

this is the density of a t_{n-2}



$\mathbb{P}\left(-t_{n-2;1-\frac{\alpha}{2}} < \frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{V}(\hat{\beta}_j)}} < t_{n-2;1-\frac{\alpha}{2}}\right) = 1-\alpha$

$\mathbb{P}\left(\hat{\beta}_j - \sqrt{\hat{V}(\hat{\beta}_j)}\, t_{n-2;1-\frac{\alpha}{2}} < \beta_j < \hat{\beta}_j + \sqrt{\hat{V}(\hat{\beta}_j)}\, t_{n-2;1-\frac{\alpha}{2}}\right) = 1-\alpha$

$\mathbb{P}(\beta_j \in \hat{C}) = 1-\alpha$ where $\hat{C} = \hat{\beta}_j \pm \sqrt{\hat{V}(\hat{\beta}_j)}\, t_{n-2;1-\frac{\alpha}{2}}$

$\hat{C}$ is a random interval. After observing the data we can compute its realization by substituting the estimators with their estimates.

We obtain $\beta_j \in \hat{\beta}_j \pm t_{n-2;1-\frac{\alpha}{2}}\sqrt{\hat{V}(\hat{\beta}_j)}$.

That is $\beta_1 \in \hat{\beta}_1 \pm t_{n-2;1-\frac{\alpha}{2}}\sqrt{s^2\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i-\bar{x})^2}\right)}$

$\beta_2 \in \hat{\beta}_2 \pm t_{n-2;1-\frac{\alpha}{2}}\sqrt{\frac{s^2}{\sum_{i=1}^n (x_i-\bar{x})^2}}$

- HYPOTHESIS TEST on β_j

$\begin{cases} H_0: \beta_j = b \\ H_1: \beta_j \neq b \end{cases}$

"under H_0" means: assuming H_0 true. We study the distribution of the test statistic assuming H_0 true, and study which values point against it.

Following the same reasoning as before, we use the TEST STATISTIC

$T_j = \frac{\hat{\beta}_j - b}{\sqrt{\hat{V}(\hat{\beta}_j)}} \overset{H_0}{\sim} t_{n-2}$ (under H_0 b is the true value of the param., so we are subtracting the true mean of $\hat{\beta}_j$)

T_j is a random variable. After observing $y_1,\dots,y_n$ we can compute the OBSERVED VALUE of THE TEST

$t^{obs}_j = \frac{\hat{\beta}_j - b}{\sqrt{\hat{V}(\hat{\beta}_j)}}$ ← we substitute the estimates: not random! it is a number

How do we define the acceptance and reject regions?

We ask: "what values of the test do we expect when H_0 is true? And what values do we expect instead when H_0 is not true (when H_1 is)?"

If H_0 true, $\beta_j = b$. If the data support this hypothesis, then the estimate $\hat{\beta}_j$ will be close to b ($\mathbb{E}[\hat{\beta}_j] = \beta_j = b$).
$\Rightarrow \hat{\beta}_j - b \approx 0 \Rightarrow t^{obs}_j \approx 0$
Hence we expect that, if H_0 is true, t^{obs} will be small (in absolute value)

If H_0 is not true, then $\beta_j \neq b$. The estimate $\hat{\beta}_j$ will be different from b
$\Rightarrow |\hat{\beta}_j - b|$ large $\Rightarrow |t^{obs}_j|$ large
Hence we expect that, under H_1, t^{obs}_j will be large (in absolute value)

The acceptance region thus will contain the values around 0 $\quad (-a, +a) = A$
The reject region will contain values far from 0 $\quad (-\infty, -a) \cup (a, +\infty) = R$
We need to define the thresholds $-a, a$

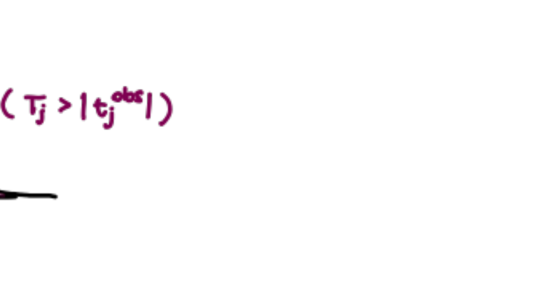
(A) fixed significance level α: $\alpha = \mathbb{P}(\text{reject } H_0 \mid H_0 \text{ true}) = \mathbb{P}_{H_0}(T_j \in R) = \mathbb{P}_{H_0}((T_j < -a) \cup (T_j > +a))$

$\mathbb{P}_{H_0}(|T_j| > t_{n-2;1-\frac{\alpha}{2}}) = \alpha$

the acceptance region is $A = (t_{n-2;\frac{\alpha}{2}},\ t_{n-2;1-\frac{\alpha}{2}})$

the reject region is $R = R_1 \cup R_2 = (-\infty;\ t_{n-2;\frac{\alpha}{2}}) \cup (t_{n-2;1-\frac{\alpha}{2}};\ +\infty)$

if $t_j^{obs} \in A \Rightarrow$ we do not reject H_0
if $t_j^{obs} \notin A \Rightarrow$ we reject H_0



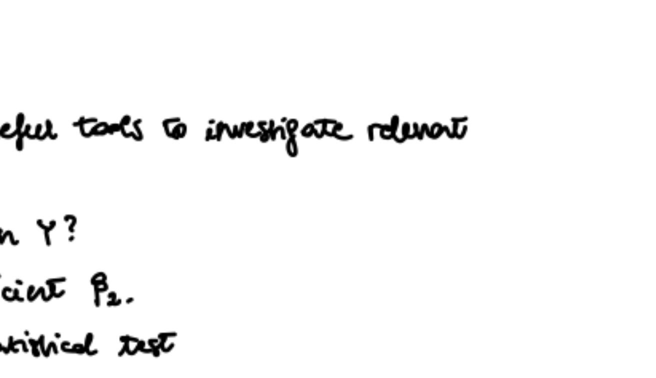
(B) p-value

it is the probability of observing "more extreme" values than t_j^{obs}

$\alpha^{obs} = 2\min\{\mathbb{P}_{H_0}(T \geq t_j^{obs});\ \mathbb{P}_{H_0}(T \leq t_j^{obs})\}$

the t distribution is symmetric, so

$\alpha^{obs} = \mathbb{P}_{H_0}(|T_j| > |t_j^{obs}|)$
$\quad = 2\cdot\mathbb{P}_{H_0}(T_j > |t_j^{obs}|)$



connection between the two types of test
- if $\alpha^{obs} < \alpha \Rightarrow$ reject H_0 at level α
- if $\alpha^{obs} > \alpha \Rightarrow$ do not reject H_0 at a level α

In practical applications, these methods are useful tools to investigate relevant applicative questions. For example:

· does the covariate x have a significant effect on Y?

The effect of x on Y is summarised by the coefficient β_2.
Hence this question can be formalized by the statistical test

$\begin{cases} H_0: \beta_2 = 0 & \rightarrow \text{no effect} \\ H_1: \beta_2 \neq 0 \end{cases}$

Indeed the model $Y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$
under H_0 becomes $Y_i = \beta_1 + \varepsilon_i$ (x has no impact on Y)

$\hookrightarrow$ this is called the "NULL MODEL"