Hypotheses testing

Bounds of the confidence interval:

$t_{c\,(\underset{81}{n}-\underset{3}{k},\,\underset{0.025}{\alpha/2})} = 1.99 \qquad \alpha = 0.05$

Ch. 2, Slide 69
Ch. 3, Slide 4

- $b_j - t_c \cdot se(b_j)$
- $b_j + t_c \cdot se(b_j)$

Essential elements of testing:

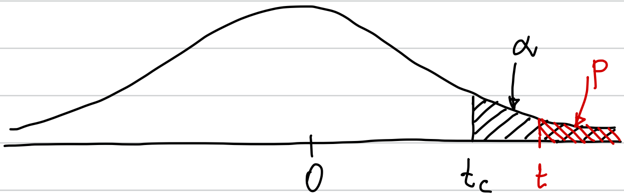
$$\boxed{t \underset{H_0}{\sim} t_{n-k}}$$

Ch. 3, Slide 5

Test statistic vs. p-value

Slide 8

An example of one-tailed t-test in the right tail of the distribution:



H_0 is rejected: $t \geq t_c$ or $p \leq \alpha$.

"Beta" coefficients or standardised regression coefficients

Slide 10

1) Standardize x_j and y:

$$x_{ji}^s = \frac{x_{ji} - \bar{x}_j}{sd(x_j)}, \quad \forall j = 1, \dots, k \qquad Var(x_j^s) = 1$$

$$y_i^s = \frac{y_i - \bar{y}}{sd(y)} \qquad Var(y^s) = 1$$

2) Estimate:

$$\widehat{y_i^s} = \underbrace{b_1}_{=0} + \underline{beta_2} \cdot x_{2i}^s + \underline{beta_3} \cdot x_{3i}^s + \dots + \underline{beta_k} \cdot x_{ki}^s$$

standardised or "beta" coefficients

3) Compare $|beta_j|$.

Higher value of $|beta_j|$ means that the variable x_j has more explanatory power in explaining the variability of y.

Statistical inference

Slide 11

- We can commit either the Type I error or the Type II error.

- Our goal is to:
  - minimize the probability of Type I error ($\alpha \leq 0.05$) and at the same time to
  - minimize the probability of Type II error or to maximize the power of the test ($1-\beta \geq 0.8$).

- Challenges:
  - there is a trade-off between α and β in practice and
  - in order to regulate β, we would have to know the actual state of nature (1 sample!).

- Reality: in practice we focus on α and at the same time we try to ensure appropriate power of the test ($1-\beta$).

- Power of the test depends on:
  1) Type of the test;
  2) Difference between the actual and the tested value (b_j vs. β_j) (+);
  3) Measurement error (−);
  4) Sample size (+).

!

- As we could commit the Type II error, we never accept H_0, only reject.
- If Ho is rejected, i.e. a significant result, we make a conclusion from H_1 (i.e. $\beta_j \neq 0$).
- If Ho is not rejected, i.e. a non-significant result, this means:
  - either no important effect (i.e. $\beta_j = 0$)
  - or too little information (could be rejected with more info. or on another sample).
- H_1 is neither accepted nor rejected, as its distribution is usually not known explicitly (not the same as that of H_0).

S14:

- c – vector of parameters of the linear combination of regression coeff. (estimates)
- r – hypothesized value (scalar) from the null hypothesis