One extra make-up class on June 25, same time

Usually we fix $\alpha = 0.05,\ 0.01$ or 0.1. This provides an upper bound for type-I error

In example 1 of Section V.1 with

$\underline{X} = (X_1, \cdots, X_N),\ X_j \sim n(\theta, 1)$

we obtained from the LRT that $\lambda(x) \le c \in (0,1) \Leftrightarrow$

$\sqrt{-2\ln(c)} \le \left| \frac{\bar{x}-\theta_0}{1/\sqrt{N}} \right| \sim n(0,1)$ for

$H_0 : \theta = \theta_0$ for a fixed θ_0

Thus, if we set $Z \sim n(0,1)$ and $P(Z \ge z_{\alpha/2}) \overset{\text{def}}{=} \frac{\alpha}{2}$

Thus, we can look for $z_{\alpha/2}$ s.t. $P(Z \ge \sqrt{-2\ln(c)}) = \frac{\alpha}{2}$

given by tables $(\Leftrightarrow P(|Z| \ge \sqrt{-2\ln(c)}) = \alpha)$

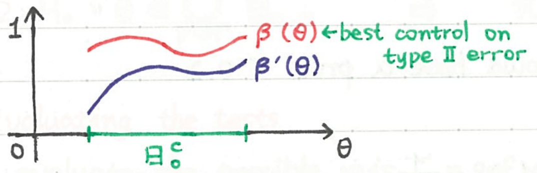
$\Leftrightarrow \sqrt{-2\ln(c)} = z_{\alpha/2}$ $\Leftrightarrow c = e^{-z_{\alpha/2}^2/2}$

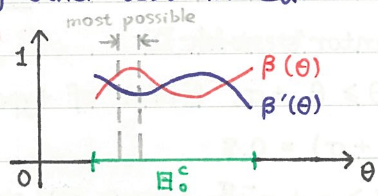
$\rightsquigarrow$ a level-α test

Def. Let C_α be the set of all level-α sets for an level-α test H_0.
A test in C_α is the uniformly most powerful (UMP) C_α-test if the corresponding power function β satisfies

$\beta(\theta) \ge \beta'(\theta)$ for any $\theta \in \Theta_0^c$ and

for the power function β' of any other test in C_α





Remark. Since different power function can cross, it is rarely possible to define the UMP C_α-test. 2 solutions:

1) Further divide Θ_0^c and take the UMP C_α-test on each part;
2) The plausible part of Θ_0^c is neglected.

V.3 p-values ($\equiv$ statistical significance)

$\rightsquigarrow$ give less arbitrariness to the value c

Aim: A p-value reports the result of a test on a more continuous scale rather than "accept H_0" or "reject H_0"

Def: Consider $H_0: \theta \in \Theta_0$ and $H_1: \theta \in \Theta_0^c$ and let $W(\underline{X})$ be a (test) statistic such that larges values of $W(\underline{X})$ give evidence that H_1 is true.

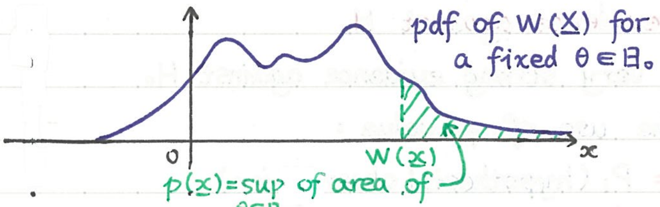
(for example $W(\underline{X}) = |X - \theta|$) For any $\underline{x}$ we set

$$p(\underline{x}) := \sup_{\theta \in \Theta_0} P_\theta(W(\underline{X}) \geq W(\underline{x})) \in [0,1]$$

Then we consider $p(\underline{X})$ and call it a p-value (for $W(\underline{X})$).

Remarks:

1) $p(\underline{X})$ is a random variable.
2) Given $\underline{x}$, $p(\underline{x})$ gives the probability of equal or more extreme values of $W(\underline{X})$, knowing that H_0 is true. $\Leftrightarrow \theta \in \Theta_0$



Example 1: $X_j \sim n(\mu, \sigma^2)$, $H_0: \mu = \mu_0$ and σ^2 arbitrary.
(μ, σ^2: unknown; μ_0: fixed)

Set $W(\underline{X}) := \dfrac{|\bar{X} - \mu_0|}{s/\sqrt{N}}$ for which large values give evidence of H_1. (recall $E(\bar{X}) = \mu$)

and which follows a $|$student t-distribution$|$ with parameter $N-1$ (indep. of σ^2)

$$p(\underline{x}) = \sup_{\sigma^2 \in \mathbb{R}} \left(P\ W(\underline{X}) \geq \frac{|\bar{x} - \mu_0|}{s/\sqrt{N}}\right) = 2P\left(T_{N-1} \geq \frac{|\bar{x} - \mu_0|}{s/\sqrt{N}}\right)$$

2: because of abs. and symmetry; T_{N-1}: student T dist.; $\frac{|\bar{x} - \mu_0|}{s/\sqrt{N}}$: found in tables

Example 2: $X_j \sim n(\mu, \sigma^2)$, $H_0: \mu \leq \mu_0 \Leftrightarrow \mu \in \Theta_0 := (-\infty, \mu_0)$

Again $W(\underline{X}) := \dfrac{\bar{X} - \mu_0}{s/\sqrt{N}}$ with the property on $W(\underline{X})$ satisfied. Then

$$\Rightarrow p(\underline{X}) = \sup_{\substack{\mu \leq \mu_0 \\ \sigma^2 \geq 0}} P(W(\underline{X}) \geq W(\underline{x})) = \sup_{\dots} P\left(\frac{\bar{X} - \mu_0 + \mu - \mu}{s/\sqrt{N}} \geq W(\underline{x})\right)$$

$$= \sup P\left(\frac{\bar{X} - \mu}{s/\sqrt{N}} \geq W(\underline{x}) + \overbrace{\boxed{\frac{\mu_0 - \mu}{S/\sqrt{N}}}}^{\geq 0}\right) = \sup P\left(\underbrace{\frac{\bar{X} - \mu}{s/\sqrt{N}}}_{\sim T_{N-1}} \geq \overbrace{\frac{\bar{x} - \mu_0}{s/\sqrt{N}}}^{W(\underline{x})}\right)$$

Remark

Here, all computations were explicit, but it is not always the case.
In general, it is not so explicit but a computer can compute $p(\underline{x})$ easily.

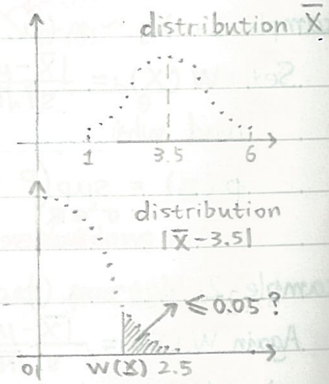
Interpretation ($p(\underline{x}) \in [0,1]$)

The p-value $p(\underline{x})$ should be interpreted in terms of repetition of the same experiment. $p(\underline{x})$ gives the prob. that the new value $W(\underline{x}')$ will be further away in the prob. distribution of $W(\underline{X})$, assuming that H_0 is correct.

p	interpretation
$p > 0.1$	No evidence against H_0 $\Rightarrow$ the data appear to be consistent with H_0
$\boxed{0.05} < p \leq 0.1$	weak evidence against H_0
$0.01 < p \leq 0.05$	moderate evidence against H_0
$p < 0.01$	strong or very strong evidence against H_0

⚠ many controversies about the use of p-value :
$P_i(\text{observation} \mid \text{hypothesis}) \neq P_i(\text{hypothesis} \mid \text{observation})$
↑ p-value ↑ what we want

Example: Roll a dice



H_0: the dice is fair (same prob of $\frac{1}{6}$ for each face)
$\underline{X} = (X_1, \cdots, X_N)$, X_j = discrete unif. dist. $\{1, 2, \cdots, 6\}$

$$W(\underline{X}) = \left| \frac{1}{N} \sum_{j=1}^{N} X_j - 3.5 \right| = |\overline{X} - 3.5|$$

large values of $W(\underline{X})$ gives evidence of H_1

If the surface of the shade is smaller than 0.05 then we reject H_0, which means that we consider that the dice is not fair.

Remark: it is a rather weak approach.
And if H_0 is rejected, it does not say anything on the dice.