V Hypothesis test

Def. A hypothesis is a statement about a population of parameter!
Two complementary hypotheses are called the null hypothesis (denoted by H_0) and the alternative hypothesis (denoted by H_1)

Example: Let $\Theta_0 \subset \Theta$ (parameter space)

$$H_0 : \theta \in \Theta_0 \text{ and } H_1 : \theta \in \underbrace{\Theta \setminus \Theta_0}_{\Theta_0^c}$$

Def. A hypothesis test is a rule which specifies for which values of $\underline{x}$ the hypothesis H_0 is accepted or rejected.

Def. The subset of the sample space for which H_0 is rejected is denoted by R and called the rejection region. More precisely:

$$R = \{\underline{x} \mid H_0 \text{ is rejected based on } \underline{x}\}$$

Example: A hypothesis test consisting in checking if the static $W(\underline{x}) \in [1,3]$ (for example $\bar{x} \in [1,3]$) in which case we accept H_0
(for example $H_0 : \underset{\text{"}\theta\text{"}}{W(\underline{X})} = 1.5$)

V.1 Finding tests

1) Likelyhood ratio test:

If $\underline{X} = (X_1, \cdots, X_N)$ with $X_j \sim f(\cdot \mid \theta)$, and recall that

$$L(\theta \mid \underline{x}) = f_{\underline{x}}(\underline{x} \mid \theta) = \prod_{j=1}^{N} f(x_j \mid \theta)$$

The Likelihood ratio test (LRT) for testing $\theta \in \Theta_0 \subset \Theta$ consists in defining

$$\lambda(\underline{x}) = \frac{\sup_{\theta \in \Theta_0} L(\theta \mid \underline{x})}{\sup_{\theta \in \Theta} L(\theta \mid \underline{x})}$$

$\cdot \in [0,1]$ is big if $\theta \in \Theta_0$ is likely

The rejection region is $R = \{\underline{x} \mid \lambda(\underline{x}) \le c\}$ for a fixed $c \in (0,1)$.
(recall $H_0 : \theta \in \Theta_0$)

Example 1: $X_j \sim n(\theta, 1)$, $H_0 : \theta = \theta_0$ (given). Then

$$\lambda(\underline{x}) = \frac{(2\pi)^{-\frac{N}{2}} \exp\left(-\sum_j (x_j \underset{-\bar{x}+\bar{x}}{-} \theta_0)^2/2\right)}{(2\pi)^{-\frac{N}{2}} \exp\left(-\sum_j (x_j - \bar{x})^2/2\right)} = \exp\left(-N(\bar{x}-\theta_0)^2/2\right)$$

$$\lambda(\underline{x}) \le c \Leftrightarrow -N(\bar{x}-\theta_0)^2/2 \le \ln(c) \Leftrightarrow |\bar{x}-\theta_0| \ge \sqrt{-\tfrac{2}{N}\ln(c)}$$

Example 2: Suppose $X_j \sim n(\mu, \overset{\text{unknown}}{\sigma^2})$; $H_0 : \mu = \mu_0$ (given)

$$\lambda(\underline{x}) = \cdots \leq c \iff \boxed{\frac{\bar{x} - \mu_0}{\underbrace{S}_{\sqrt{\text{sample variance}}}/\sqrt{N}}} \geq c \quad \text{(computation done in Appendix 7)}$$

$\hookrightarrow$ follows a student t distribution

2) Bayesian test

Prior $\overset{\text{prob. distribution for } \theta}{\text{distribution}}$ π for θ, and we want to compute π_{post} for θ.

Then we are going to accept $H_0 : \theta \in \Theta_0$ if

$$P_{post}(\theta \in \Theta_0 \mid \underline{x}) = \int_{\Theta_0} \pi_{post}(\theta \mid \underline{x})\, d\theta \geq c$$

Example: $X_j \sim n(\theta, \sigma^2)$; $\pi \sim n(\mu, \tau^2)$ with σ^2, μ, τ^2 known.

$\uparrow$ not an accident

$H_0 : \theta \leq \theta_0$ fixed

$$\pi_{post}(\cdot \mid \underline{x}) = \pi_{post}(\cdot \mid \bar{x}) \overset{\text{Ex 7.22}}{=} n\left(\frac{N\tau^2\bar{x} + \sigma^2\mu}{N\tau^2 + \sigma^2}, \frac{\sigma^2\tau^2}{N\tau^2 + \sigma^2}\right)$$

If we choose $c = \frac{1}{2}$, we get by symmetry

$$P_{post}(\theta \leq \theta_0 \mid \underline{x}) \geq \frac{1}{2} \iff \frac{N\tau^2\bar{x} + \sigma^2\mu}{N\tau^2 + \sigma^2} \leq \theta_0$$

$$\iff \bar{x} \leq \theta_0 + \frac{\sigma^2(\theta_0 - \mu)}{N\tau^2}$$

3) Union-intersection and intersection-union tests

Case 1: $H_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_{0,\gamma}$ $\Rightarrow$ $R = \bigcup_{\gamma \in \Gamma} R_\gamma$ (U.-I. test)

$\Gamma \to$ family of condition

Case 2: $H_0 : \theta \in \bigcup_{\gamma \in \Gamma} \Theta_{0,\gamma}$ $\Rightarrow$ $R = \bigcap_{\gamma \in \Gamma} R_\gamma$ (I.-U. test)

V.2 Evaluating the tests

Idea: evaluate the possible mistakes of a hypothesis test:

		Decision (on $\underline{x}$)	
		accept H_0	reject H_0
"Truth"	H_0	♡	Type-I error
	H_1	Type-II error	♡

If $\theta \in \Theta_0$ and R denotes the rejection region (for H_0),

then an error of type I takes place if $\underline{x} \in R$. Its probability is $P_\theta(\underline{x} \in R)$

Conversely if $\theta \in \Theta_0^c$, the prob. of a type II error is $1 - P_\theta(\underline{x} \in R)$

Def. The power function of a hypothesis test (c) with the rejection region R is

$\beta : \theta \longmapsto P_\theta(\underline{x} \in R)$

Ideal situation: if $\theta \in \Theta_0$, $\beta(\theta) \sim 0$

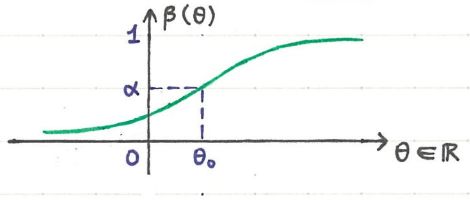
if $\theta \in \Theta_0^c$, $\beta(\theta) \sim 1$

Example: $X_j \sim n(\theta, \sigma^2)$; $H_0 : \theta \leq \theta_0 \Leftrightarrow \Theta_0 = (-\infty, \theta_0]$
known

The hypothesis H_0 is rejected if $\frac{\bar{X} - \theta_0}{\sigma/\sqrt{N}} \geq c$. Thus
fixed later on

$$\beta(\theta) = P_\theta\left(\frac{\bar{X}-\theta_0}{\sigma/\sqrt{N}} \geq c\right) = P_\theta\left(\frac{\bar{X}-\theta}{\sigma/\sqrt{N}} \geq c + \frac{\theta_0 - \theta}{\sigma/\sqrt{N}}\right)$$

$$= P_\theta\left(Z \geq c + \frac{\theta_0 - \theta}{\sigma/\sqrt{N}}\right) \text{ with } Z \sim n(0,1)$$



with $\alpha = P(Z \geq c)$

Experimentor's wish 1:

Error type I with prob ≤ 0.1

$\Leftrightarrow \beta(\theta) \leq 0.1 \ \forall \theta \leq \theta_0$ $\Leftrightarrow \beta(\theta_0) = 0.1$ since β is increasing

$\Leftrightarrow P_\theta(Z \geq c) = 0.1$ $\overset{\text{table}}{\Leftrightarrow}$ $\underline{c = 1.28}$

Experimentor's wish 2:

For $\theta \geq \theta_0 + \sigma$, Error of type II should have a prob. ≤ 0.2

$\Leftrightarrow \beta(\theta_0 + \sigma) = 0.8$

$\Leftrightarrow P_\theta\left(Z \geq c + \frac{-\sigma}{\sigma/\sqrt{N}}\right) = 0.8 \overset{\text{tables}}{\Leftrightarrow} -0.84 \geq 1.28 - \sqrt{N}$

$\Leftrightarrow \underline{N \geq 5}$

⚠ It is not always possible to adjust both errors.

We usually concentrate on error of type I.

Def. For any $\alpha \in [0,1]$, a test of power function β

is a size α-test if $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$, and

is a level α-test if $\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$.

} way to measure the prob. of type-I error.