The Elements of a Statistical Test

1. Null hypothesis, H_0
2. Alternative hypothesis, H_a
3. Test statistic
4. Rejection region

DEFINITION 10.1 A *type I error* is made if H_0 is rejected when H_0 is true. The *probability of a type I error* is denoted by α. The value of α is called the *level* of the test.

A *type II error* is made if H_0 is accepted when H_a is true. The *probability of a type II error* is denoted by β.

Large-Sample α-Level Hypothesis Tests

$$H_0: \theta = \theta_0.$$

$$H_a: \begin{cases} \theta > \theta_0 & \text{(upper-tail alternative).} \\ \theta < \theta_0 & \text{(lower-tail alternative).} \\ \theta \neq \theta_0 & \text{(two-tailed alternative).} \end{cases}$$

$$\text{Test statistic: } Z = \frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}}.$$

$$\text{Rejection region: } \begin{cases} \{z > z_\alpha\} & \text{(upper-tail RR).} \\ \{z < -z_\alpha\} & \text{(lower-tail RR).} \\ \{|z| > z_{\alpha/2}\} & \text{(two-tailed RR).} \end{cases}$$

the probability β of a type II error is

$$\begin{aligned} \beta &= P(\hat{\theta} \text{ is not in RR when } H_a \text{ is true}) \\ &= P(\hat{\theta} \leq k \text{ when } \theta = \theta_a) = P\left(\frac{\hat{\theta} - \theta_a}{\sigma_{\hat{\theta}}} \leq \frac{k - \theta_a}{\sigma_{\hat{\theta}}} \text{ when } \theta = \theta_a\right). \end{aligned}$$

Sample Size for an Upper-Tail α-Level Test

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2}$$

DEFINITION 10.2 If W is a test statistic, the *p-value*, or *attained significance level*, is the smallest level of significance α for which the observed data indicate that the null hypothesis should be rejected.

RR: $\{w \le k\}$—the p-value associated with an observed value w_0 of W is given by

$$p\text{-value} = P(W \le w_0, \text{ when } H_0 \text{ is true}).$$

Analogously, if we were to reject H_0 in favor of H_a for large values of W—say, RR: $\{w \ge k\}$—the p-value associated with the observed value w_0 is

$$p\text{-value} = P(W \ge w_0, \text{ when } H_0 \text{ is true}).$$

A Small-Sample Test for μ

Assumptions: $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from a normal distribution with $E(Y_i) = \mu$.

$H_0: \mu = \mu_0$.

$$H_a: \begin{cases} \mu > \mu_0 & \text{(upper-tail alternative).} \\ \mu < \mu_0 & \text{(lower-tail alternative).} \\ \mu \ne \mu_0 & \text{(two-tailed alternative).} \end{cases}$$

Test statistic: $T = \dfrac{\overline{Y} - \mu_0}{S/\sqrt{n}}$.

$$\text{Rejection region: } \begin{cases} t > t_\alpha & \text{(upper-tail RR).} \\ t < -t_\alpha & \text{(lower-tail RR).} \\ |t| > t_{\alpha/2} & \text{(two-tailed RR).} \end{cases}$$

(See Table 5, Appendix 3, for values of t_α, with $\nu = n - 1$ df.)

Small-Sample Tests for Comparing Two Population Means

Assumptions: Independent samples from normal distributions with $\sigma_1^2 = \sigma_2^2$.

$H_0: \mu_1 - \mu_2 = D_0$.

$$H_a: \begin{cases} \mu_1 - \mu_2 > D_0 & \text{(upper-tail alternative).} \\ \mu_1 - \mu_2 < D_0 & \text{(lower-tail alternative).} \\ \mu_1 - \mu_2 \ne D_0 & \text{(two-tailed alternative).} \end{cases}$$

Test statistic: $T = \dfrac{\overline{Y}_1 - \overline{Y}_2 - D_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, where $S_p = \sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$.

$$\text{Rejection region: } \begin{cases} t > t_\alpha & \text{(upper-tail RR).} \\ t < -t_\alpha & \text{(lower-tail RR).} \\ |t| > t_{\alpha/2} & \text{(two-tailed RR).} \end{cases}$$

Here, $P(T > t_\alpha) = \alpha$ and degrees of freedom $\nu = n_1 + n_2 - 2$. (See Table 5, Appendix 3.)

Test of Hypotheses Concerning a Population Variance ← equivalent for σ

Assumptions: $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from a normal distribution with

$$E(Y_i) = \mu \quad \text{and} \quad V(Y_i) = \sigma^2.$$

$H_0 : \sigma^2 = \sigma_0^2$

$H_0: \sigma = \sigma_0 \iff H_0: \sigma^2 = \sigma_0^2$

$$H_a : \begin{cases} \sigma^2 > \sigma_0^2 & \text{(upper-tail alternative).} \\ \sigma^2 < \sigma_0^2 & \text{(lower-tail alternative).} \\ \sigma^2 \neq \sigma_0^2 & \text{(two-tailed alternative).} \end{cases}$$

Test statistic: $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$.

$$\text{Rejection region:} \begin{cases} \chi^2 > \chi_\alpha^2 & \text{(upper-tail RR).} \\ \chi^2 < \chi_{1-\alpha}^2 & \text{(lower-tail RR).} \\ \chi^2 > \chi_{\alpha/2}^2 \text{ or } \chi^2 < \chi_{1-\alpha/2}^2 & \text{(two-tailed RR).} \end{cases}$$

Notice that χ_α^2 is chosen so that, for $\nu = n - 1$ df, $P(\chi^2 > \chi_\alpha^2) = \alpha$. (See Table 6, Appendix 3.)

DEFINITION **10.3**

Suppose that W is the test statistic and RR is the rejection region for a test of a hypothesis involving the value of a parameter θ. Then the *power* of the test, denoted by power(θ), is the probability that the test will lead to rejection of H_0 when the actual parameter value is θ. That is,

$$\text{power}(\theta) = P(W \text{ in RR when the parameter value is } \theta).$$

Power $(\theta_0) = \alpha = P(W \text{ in RR when } \theta = \theta_0)$

Relationship Between Power and β

If θ_a is a value of θ in the alternative hypothesis H_a, then

$$\text{power}(\theta_a) = 1 - \beta(\theta_a).$$

DEFINITION **10.4**

If a random sample is taken from a distribution with parameter θ, a hypothesis is said to be a *simple hypothesis* if that hypothesis *uniquely specifies* the distribution of the population from which the sample is taken. Any hypothesis that is not a simple hypothesis is called a *composite hypothesis*.

THEOREM **10.1**

The Neyman–Pearson Lemma Suppose that we wish to test the simple null hypothesis $H_0 : \theta = \theta_0$ versus the simple alternative hypothesis $H_a : \theta = \theta_a$, based on a random sample $Y_1, Y_2, \ldots, Y_n$ from a distribution with parameter θ. Let $L(\theta)$ denote the likelihood of the sample when the value of the parameter is θ. Then, for a given α, the test that maximizes the power at θ_a has a rejection region, RR, determined by

$$\frac{L(\theta_0)}{L(\theta_a)} < k.$$

The value of k is chosen so that the test has the desired value for α. Such a test is a most powerful α-level test for H_0 versus H_a.

A Likelihood Ratio Test

Define λ by

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{\max\limits_{\Theta \in \Omega_0} L(\Theta)}{\max\limits_{\Theta \in \Omega} L(\Theta)}.$$

A likelihood ratio test of $H_0 : \Theta \in \Omega_0$ versus $H_a : \Theta \in \Omega_a$ employs λ as a test statistic, and the rejection region is determined by $\lambda \le k$.