Power analysis

Miaoyan Wang

Department of Statistics
UW Madison

Review of one-sample testing

- i.i.d. sample: $Y_1, Y_2, \ldots, Y_n \sim N(\mu, \sigma^2)$
- Parameter of interest: μ or σ.
- H_0 versus H_A
- Testing $H_0 : \mu = \mu_0$
  - Normal population σ^2 known:

$$Z = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

  - Normal population σ^2 unknown:

$$T = \frac{\bar{Y} - \mu_0}{\hat{\sigma}/\sqrt{n}} \sim T_{n-1}, \quad \text{where } \hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2},$$

- Testing for σ^2: $H_0 : \sigma^2 = \sigma_0^2$

$$V^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$$

- p-value and conclusion

Example: $H_0 : \mu = 60$ vs. $H_A : \mu \neq 60$

- Consider testing
$$H_0 : \mu = 60 \text{ vs. } H_A : \mu \neq 60.$$
- Suppose that an i.i.d. sample of size $n = 16$ is available and that the population distribution is $N(\mu, 36)$.
- Under H_0, what is the distribution of the sample mean $\bar{Y}$?
$\bar{Y} \sim N(\mu_0 = 60, \frac{\sigma^2}{n} = \frac{36}{16} = (1.5)^2)$
- Suppose $\alpha = 0.05$. That is, we reject H_0 if the p-value is less than or equal to 0.05.
- What values of $\bar{Y}$ would lead to a rejection of H_0?
- Can we answer this question before even having the data?

Rejection Region

- Note $z_{0.025} = 1.96$ (from R or Z-table). We would reject H_0 if the observed $\bar{y}$ is more than 1.96 standard deviation (of the sample mean!) away from the hypothesized mean $\mu = 60$.

- We would reject H_0 if the observed $\bar{y}$ is less than 57.06 or more than 62.94. [Why?]

$$
\begin{aligned}
0.05 &= P(Z \leq -1.96) + P(Z \geq 1.96) \\
&= P\left(\frac{\bar{Y} - 60}{1.5} \leq -1.96\right) + P\left(\frac{\bar{Y} - 60}{1.5} \geq 1.96\right) \\
&= P(\bar{Y} \leq 60 - 1.96 \times 1.5) + P(\bar{Y} \geq 60 + 1.96 \times 1.5) \\
&= P(\bar{Y} \leq 57.06) + P(\bar{Y} \geq 62.94)
\end{aligned}
$$

Type I and II Error

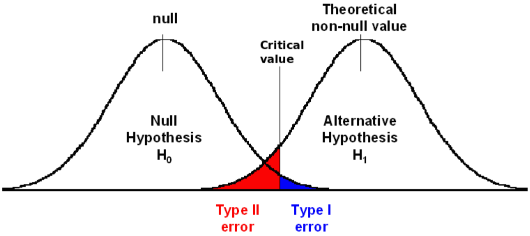
- For a given rejection region, the rule is to reject H_0 if the observed $\bar{y}$ falls in the rejection region and do not reject H_0 otherwise.
- Because $\bar{Y}$ is random, it is possible to make two types of errors.
  - A **Type I error** occurs if we reject H_0 when H_0 is true.
  - A **Type II error** occurs if we accept H_0 when H_0 is false.
- In our example, what is the probability of Type I error?

$$P(\text{Reject } H_0 | H_0) = P(\bar{Y} \leq 57.06 \text{ or } \bar{Y} \geq 62.94 | \mu = 60)$$

Types of Error and Statistical Power

There are four possible outcomes that could be reached as a result of the null hypothesis being either true or false and the decision being either “fail to reject” or “reject”.

Reality	Our Decision	
	H_0	H_a
H_0	$\surd$	Type I Error
H_a	Type II Error	$\surd$



Types of Error and Statistical Power

Reality	Our Decision: H_0	Our Decision: H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

- The **significance level** α of any fixed level test is the probability of a Type I error.
- The **power** of a fixed level test against a particular alternative is $1 - \beta$ for that alternative.
- In practice, we first choose an α and consider only tests with probability of Type I error no greater than α. Then we select one that makes the probability of Type II error as small as possible (i.e. the most powerful possible test).

Power

- For a given rejection rule and for any given value of μ, the **power** is the probability of rejecting H_0 given the value of μ.
- In the example above, what is the power for $\mu = 60$?

$$P(\text{Reject } H_0 | \mu = 60) = P(\bar{Y} \leq 57.06 \text{ or } \bar{Y} \geq 62.94 | \mu = 60)$$

- What is the power for $\mu = 62$?

$$P(\text{Reject } H_0 | \mu = 62) = P(\bar{Y} \leq 57.06 \text{ or } \bar{Y} \geq 62.94 | \mu = 62)$$

$$= P(Z \leq -3.29) + P(Z \geq 0.63) = 0.0005 + 0.2643 = 0.2648$$

- What is the power for $\mu = 64$?

$$P(\text{Reject } H_0 | \mu = 64) = P(\bar{Y} \leq 57.06 \text{ or } \bar{Y} \geq 62.94 | \mu = 64)$$

$$= P(Z \leq -6.94) + P(Z \geq -0.71) = 0 + (1 - 0.2389) = 0.7611$$