Beta and Power

Data Science for Quality Management: Two Sample Hypothesis Testing

with **Wendy Martin**

Learning objective:

Interpret the relationship between Beta and Power

Beta and Power

- Power is the probability of correctly rejecting a False H_0
- Correct decision when H_0 is false
- Power is designated as $1 - \beta$

Beta and Power

- Power is used in determining how well a test is working and likely to detect a true effect (or difference)

Beta and Power

- Affected by
  - True value of the population parameter
  - Significance level, α
  - Standard deviation, s (or σ)
  - Sample size, n

Strategies for Considering Power in Experiments

Strategy 1

- Fix the probability of committing a Type I error (α).

Strategies for Considering Power in Experiments

Strategy 1

- Select a sample size large enough so that β is acceptably small, and testing is not too expensive or time consuming to conduct.

Strategies for Considering Power in Experiments

Strategy 2

- Consider the Null and Research Hypotheses and select the α and β pair which best represents your wishes related to the research,

Calculating Beta and Power for Means

Strategy 2

- Then, calculate the sample size required to maintain the selected risk levels

Example

Assumptions

- The Central Limit Theorem is applicable
- The RSD's employed will be approximately normal

Example

An engineer has been studying the effects of modifying a component. They want to know if the change significantly reduces an output measure called response time, measured in milliseconds.

Example

Historical data indicates that the response time average, μ, is 20 ms with a standard deviation of 3.5 ms.

Example

Further, let us assume that a reduction in the mean of at least 4 ms is necessary before the modification becomes cost-effective. This is referred to as the effect size (Δ) of an experiment.

Example

In this case, the hypotheses tested might be as follows.

$$H_0: \mu \geq 20$$
$$H_1: \mu < 20$$

Example

Based on the original data, we recall that the process μ has been 20; therefore,

if H_0 is true, $\mu \geq 20$ and $\sigma = 3.5$

if H_1 is true, $\mu \leq 16$ and $\sigma = 3.5$

Given a Δ of 4 ms

Example

Assume further that α was selected at a 0.05 level

Note also that a one-tailed test has been employed for the hypotheses to be tested, which corresponds to a z of –1.645.

Example

We will assume that $n = 9$.

What would be the power to detect a difference of 4 units?

Calculating β and Power for Means

Step 1:

Determine the critical value in the H_0 RSD of means corresponding to the z value for the given value of α

Calculating β and Power for Means

$$z_{\bar{X}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$-1.645 = \frac{\bar{X} - 20}{\frac{3.5}{\sqrt{9}}} = \frac{\bar{X} - 20}{1.167}$$

$$\bar{X} = 18.081$$

Calculating β and Power for Means

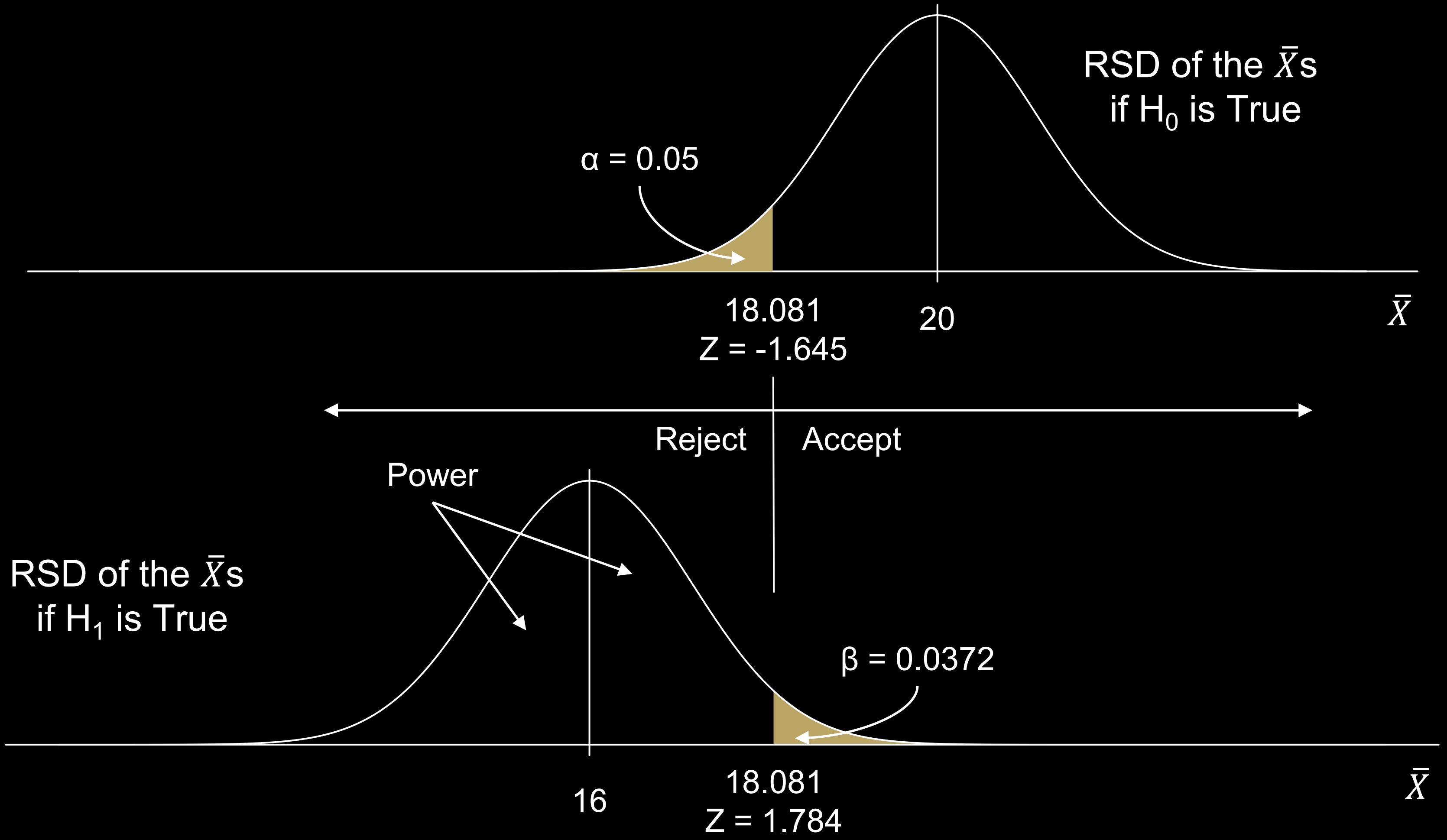
Step 2:

Calculate the z value and area corresponding to the calculated on the "H_1 is true" curve

Calculating β and Power for Means

$$z = \frac{18.081 - 16}{1.167} = 1.784$$

$$z = 1.784$$
$$\beta = 0.0372$$



RSD of the $\bar{X}$s
if H_0 is True
α = 0.05
18.081
Z = -1.645
20
$\bar{X}$
Reject
Accept
Power
RSD of the $\bar{X}$s
if H_1 is True
β = 0.0372
16
18.081
Z = 1.784
$\bar{X}$

Summary Observations

- The larger the value of $\mu_0 – \mu_1$ (Δ delta), the larger power $(1 – \beta)$ will become
- Generally, both α and β should be small. In industry, studies planned without an initial regard for β generally result in low power or high β values

Summary Observations

- It is not possible to commit Type I and Type II errors at the same time
- Had a two-tailed test been employed, the power of the test would have been the sum of the two areas falling beyond α on the H_1 distribution
- Increasing α will generally reduce β

Summary Observations

- Increasing n will generally increase the power of the test
- Increasing the power of the test can be accomplished by reducing the standard error through design modifications (for example, matched groups and stratified sampling)

Sources

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982