PROBABILITY & STATISTICS FOR LIFE SCIENCES

Alpha, Beta, and Power

Exam cram • Zan Ahmad

Read the condition first

In $\mathbb{P}(\text{decision} \mid \text{assumed truth})$, the words after the bar select the population model. Imagine many samples from *that model*; ask what fraction produce the decision before the bar. This is not the probability that a hypothesis is true.

Four quantities to keep separate

Type I error: reject a true null. The significance level controls false alarms:

$$\alpha = \mathbb{P}(\text{reject } H_0 \mid H_0 \text{ true}).$$

***p*-value: evidence from the observed result.** Under H_0, find the probability of a statistic at least as extreme as the observed one, in the direction(s) of H_a. Reject when $p \leq \alpha$.

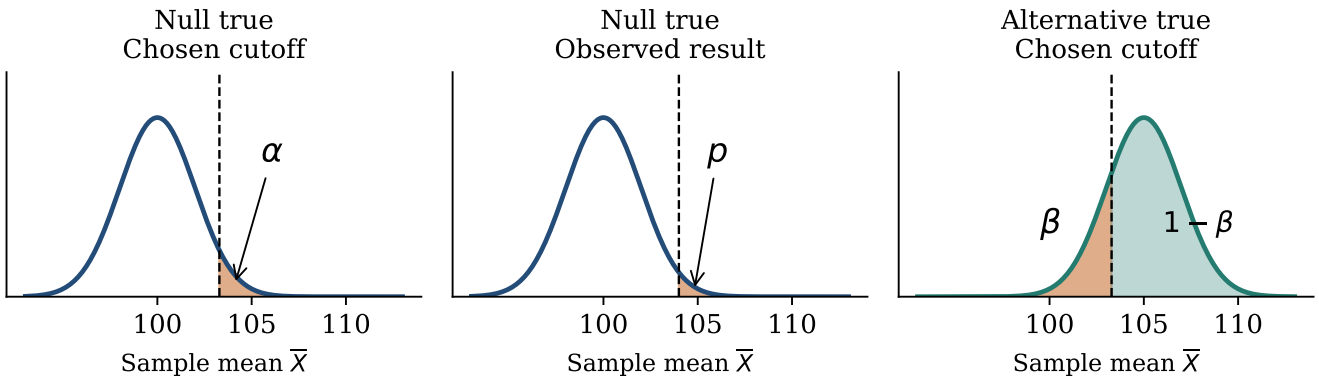
Type II error: fail to reject a false null. For a *specified* alternative value μ_1,

$$\beta(\mu_1, \alpha) = \mathbb{P}(\text{fail to reject } H_0 \mid \mu = \mu_1).$$

Power: detect that effect. Under the *same alternative*,

$$1 - \beta = \mathbb{P}(\text{reject } H_0 \mid \mu = \mu_1).$$

Three shaded areas tell three different stories



Left: α uses the *chosen cutoff* under the null. **Middle:** p uses the *observed statistic* under the null. **Right:** β and power use the *chosen cutoff* under a specified alternative. Here a right-tailed test is illustrated; the sample mean has SD 2, with null mean 100 and alternative mean 105.

Do not reverse the conditioning

Neither p nor α is the probability that H_0 is true or that this particular decision is wrong. Also, $\alpha + \beta$ need not equal 1, and $1 - p$ is not power. **Failing to reject does not prove no effect.**

Turn a question into a probability

1. **Fix the rule.** Choose the tail(s) from H_a and use α to obtain the rejection cutoff(s). Do not change this rule when computing β.
2. **Identify what is being asked.** For a p-value, use the *null distribution* and observed statistic. For β or power, specify the alternative effect and use its sampling distribution.
3. **Shade before calculating.** Under the alternative, the nonrejection region is β; the rejection region is power.

Course notation. z_α and $t_{\mathrm{df},\alpha}$ have right-tail area α. A two-sided test uses $\alpha/2$ in each tail. Use the appropriate null N or T distribution for the p-value and critical values.

Beta when the sample mean is normal under the alternative

Write $\mathrm{SE} = \sigma/\sqrt{n}$ for the *known* standard error and Φ for the standard-normal CDF. Under the specified alternative, center at μ_1:

$$\begin{aligned}
\text{Reject if } \overline{X} > c: &\quad \beta = \Phi\left(\frac{c-\mu_1}{\mathrm{SE}}\right), \\
\text{Reject if } \overline{X} < c: &\quad \beta = 1 - \Phi\left(\frac{c-\mu_1}{\mathrm{SE}}\right), \\
\text{Reject outside } [L,U]: &\quad \beta = \Phi\left(\frac{U-\mu_1}{\mathrm{SE}}\right) - \Phi\left(\frac{L-\mu_1}{\mathrm{SE}}\right).
\end{aligned}$$

Then power is $1-\beta$. These formulas are exact for a normal sample mean with known σ, not automatically for an unknown-σ t test. Exact t-test power uses a noncentral t distribution.

One quick example

Test $H_0 : \mu = 100$ versus $H_a : \mu > 100$, with independent normal observations, known $\sigma = 10$, $n = 25$, and $\alpha = 0.05$. Then $\mathrm{SE} = 2$ and the rule rejects if $\overline{X} > 103.29$.

- **Observed $\bar{x} = 104$:** $z = 2$, so $p = 1 - \Phi(2) \approx 0.0228$. Reject H_0.
- **Assumed true $\mu_1 = 105$:** $\beta = \Phi((103.29 - 105)/2) \approx 0.196$. Power is about 0.804. These are repeated-study probabilities, not probabilities about whether the null is true in the observed study.

Fast checks for exam problems

Smaller effect? More difficult to detect: larger β, lower power. **Larger n or smaller variability?** Usually higher power at the same α. **Larger α?** More power, but more false alarms, with other inputs fixed.

Try it: Keep the example's rule but assume $\mu_1 = 103$. Which probability gives β? *Answer:* $\Phi((103.29 - 103)/2) = \Phi(0.145) \approx 0.558$; power ≈ 0.442. The cutoff stays fixed; the assumed center changes.