Math 373  Lecture 19

Type I, II Errors, β and Power

There are two types of error:

- **Type I error**: Falsely rejecting \( H_0 \). (\( H_0 \) is true but rejected.)
- **Type II error**: Falsely accepting \( H_0 \). (\( H_0 \) false but accepted.)

**Theorem.** The probability of falsely rejecting \( H_0 \) is the level of significance \( \alpha \).

**Proof.** Suppose \( H_0 \) is falsely rejected. Then (a) \( H_0 \) is true and (b) the statistic is in the rejection region. The probability of this is \( \alpha \).

**Definition.** Let \( \beta = P(\text{falsely accepting } H_0) = P(\text{accepting } H_0 \text{ when it is false}) \). The power of the test is \( P(\text{correctly rejecting } H_0) = P(\text{accepting } H_a \text{ when it is true}) \).

**Theorem.** The power of the test is \( 1 - \beta \).

**Proof.** The power \( = P(\text{correctly rejecting } H_0) = P(\text{rejecting } H_0 \text{ when it is false}) = 1 - P(\text{accepting } H_0 \text{ when it is false}) = 1 - \beta \).

Since \( \alpha \) and \( \beta \) are the probabilities of type I and type II errors, we want both to be as low as possible. We can choose the level of significance \( \alpha \) to be as low as we wish, but reducing the probability \( \alpha \) of making a type I error, increases the probability \( \beta \) of making a type II error. The only way to decrease both types of error is to increase the sample size.

Suppose the null hypothesis is \( \mu = \mu_0 \) where \( \mu_0 \) is the **null value**. The alternate hypothesis is \( H_a: \mu \neq \mu_0 \). This says what \( \mu \) is not but doesn’t give us enough information to calculate \( \beta \) or the power. But suppose we believe \( \mu \) to be a specific **alternate value** \( \mu_a \). Then we can replace the alternate hypothesis \( \mu \neq \mu_0 \) with the more specific alternate hypothesis \( H_a: \mu = \mu_a \).

Given a specific alternate value \( \mu_a \), we can calculate \( \beta \). Suppose the acceptance region is \( \mu_0 \pm z_{\alpha/2}\text{SE} = [a_1, a_2] \). By definition, \( \beta \) is the probability of accepting \( H_0 \) when it is false. Accepting \( H_0 \) when it is false means:

- \( H_0 \) is true. (Hence the heavy curve is the true distribution.)
- The statistic \( \mu \) is in the acceptance region (grey area).

Thus \( \beta \) is the area under the solid heavy bell curve between \( a_1 \) and \( a_2 \) (Cross hatched area.)

If the null hypothesis had been true (dotted line), then the probability of the acceptance region would have been the shaded area with probability \( 1 - \alpha \). But under the alternate hypothesis \( \mu = \mu_a \), the probability of the acceptance region is the area pictured with diagonal rulings. Under the alternate hypothesis, the normal curve (heavy solid curve) till has std. dev. = SE but the mean \( \mu \) is \( \mu_a \) instead of \( \mu_0 \).

- Suppose the null value is \( \mu_0 = 50 \) and the alternate value is \( \mu_a = 52 \). Suppose SE = 1 and the significance level is \( \alpha = 5\% \). Find \( \beta \) and the power of the test.

Find or calculate the following:

- \( z_{\alpha/2} \) and SE, SE is given here but not in the homework.
- acceptance region \([a_1, a_2]\), use the null hypothesis \( \mu = \mu_0 = 50 \).
- \( z \)-scores of \( a_1 \) and \( a_2 \), use alternate hypothesis that \( \mu = \mu_a = 52 \).
- the probability \( \beta = P([a_1, a_2]) \).

\[ z_{\alpha/2} = 1.96, \text{ SE } = 1. \] In your homework, you calculate SE from \( s \).

\[ [a_1, a_2] = [\mu_0 \pm z_{\alpha/2}\text{SE}] = [50 \pm (1.96)(1)] = [48.04, 51.96] \]

\( z \)-score of 48.04* = \( (\mu - \mu_a)/\text{SE} = (48.04 - 52)/1 = -3.96 \)

\( z \)-score of 51.96 = \( (\mu - \mu_a)/\text{SE} = (51.96 - 52)/1 = -0.4 \)

Since the tail left of -3.96 has approximately 0 area, we can replace it with \( -\infty \).

\( \beta = P(z \in [-3.96, -0.4]) \approx P(z \in (-\infty, -0.4]) = 0.484 \).

Thus \( \beta = 0.48 \).

The power of the test is \( 1 - \beta = 0.52 \).

*The acceptance region \([a_1, a_2]\) is calculated under the assumption of the null hypothesis and uses the null mean \( \mu_0 \). But the probability of \([a_1, a_2]\) is calculated with the alternate hypothesis. Thus the \( z \)-scores of \( a_1 \), \( a_2 \) are computed using the alternate value \( \mu_a \) as the mean.

Type II errors are considered less serious than type I errors but still, \( \beta = 48\% \) is an unacceptably high error rate. \( \beta \) is high since the difference between \( \mu_a \) and \( \mu_0 \) is too close to the margin of error (1.96×1). To adequately distinguish between \( \mu_0 \) and \( \mu_a \), we must increase the sample size and reduce the margin of error so that \( \mu_0 \) and \( \mu_a \) can be distinguished with a reasonably small (say 10%) probability of type II error.

In general, we don’t know the alternate value \( \mu_a \). But if we can estimate how far it might be from \( \mu_0 \), then we can estimate \( \beta \) as above.

Type II errors (falsely accepting \( H_0 \)) are less serious since \( H_0 \) is the status quo. No change is needed if \( H_0 \) is accepted. A type I error (falsely accepting \( H_a \)) is more serious. It means abandoning the status quo for a false assumption. There may be a considerable investment behind the status quo and change may be expensive. In our legal system, a type II error (freeing a guilty man) is less serious than a type I error (convicting an innocent man).