

QUESTIONS (Test Bank)

Chapter 1 Review of Estimation and Hypothesis Tests

1. Explain the following.

- | | | |
|-------------------------|---------------------------|-----------------|
| (a) estimator, estimate | (b) sampling distribution | |
| (c) Type I error | (d) significance level | |
| (e) p -value | (f) rejection region | (g) power |
| (h) unbiasedness | (i) efficiency | (j) consistency |

2. Consider a random sample $\{Y_1, \dots, Y_n\}$ from a population with mean μ and variance σ^2 . That is, $E(Y_i) = \mu$, $Var(Y_i) = \sigma^2$ and $Cov(Y_i, Y_j) = 0$ if $i \neq j$. We are going to use the sample mean $\bar{Y} = (Y_1 + \dots + Y_n)/n$ to estimate and test about the population mean.

(a) What is the expected value of $\bar{Y}$?

(b) What is the variance of $\bar{Y}$?

3. Suppose we test $H_0: \mu = 15$ vs. $H_0: \mu \neq 15$. If we cannot reject H_0 at the 5% significance level, can we conclude that the true population mean μ is equal to 15?

4. The power of a test is defined as $\Pr(H_0 \text{ is rejected} \mid H_1 \text{ is true})$, the probability of correctly rejecting the null hypothesis when the alternative hypothesis is true. The powerful test is the one which maximizes the power. For the following hypotheses, what is the rejection region for the powerful test?

$$H_0: \mu = 28 \text{ vs } H_1: \mu < 28$$

5. Explain the following using sampling distributions.

- (a) If an estimator is unbiased, is it always consistent?
- (b) If an estimator is biased, is it consistent?

(c) Is it possible that an estimator is biased but consistent?

6. According the Central Limit Theorem, the sample mean ($\bar{Y}$) is known to follow a normal

distribution for sufficiently large sample sizes, i.e.,
$$\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Explain whether $\bar{Y}$ is unbiased and consistent.

ANSWERS (Test Bank)

Chapter 1 Review of Estimation and Hypothesis Tests

1.

(a) An *estimator* is a formula that assigns a value to each possible outcome of a sample. The value assigned by an estimator is called an *estimate*. For example, $\bar{Y} = (Y_1 + \dots + Y_n)/n$ is an estimator for μ and the sample mean 26.8 calculated using the 49 observations is an estimate.

(b) *sampling distribution*

This is the distribution of all possible values of an estimator. The estimator yields different estimates depending on which sample is drawn from the population.

(c) *Type I error*

This is an error that the null hypothesis is rejected when the null hypothesis is true.

(d) *significance level*

This is the probability of making a type I error. In other words, this is the probability that an observed value of the test statistic falls in the rejection region when the null hypothesis is true.

(e) *p-value*

This is the smallest significance level at which H_0 can be rejected.

(f) *rejection region*

This is a region (or an interval) whose values are least likely to occur if the null hypothesis is correct. Therefore, if an observed estimate falls in the rejection region, we reject H_0 ,

(g) *power*

The power is the probability that the null hypothesis is rejected when it is not true, $\Pr(H_0 \text{ is rejected} \mid H_0 \text{ is NOT true})$. Thus a powerful test is more likely to reject the null when it is not true.

(h) *unbiasedness*

An estimator is said to be unbiased if its expected value is equal to the parameter to be estimated; e.g., $E(\bar{Y}) = \mu$.

(i) *efficiency*

An estimator is said to be efficient if it has the smallest variance among unbiased estimators.

(j) *consistency*

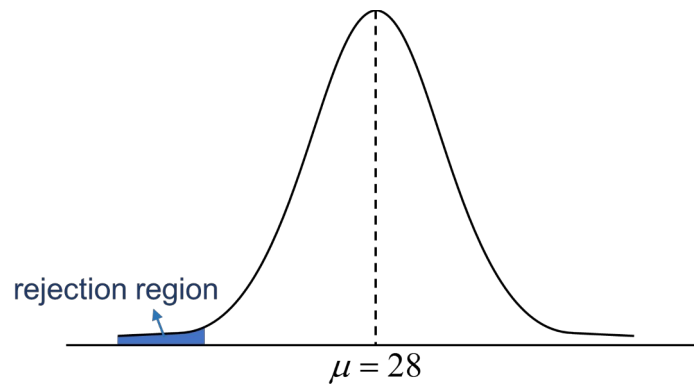
$\text{plim } \bar{Y} = \mu$. It implies that estimates $\bar{Y}$ obtained from infinitely many observations will be equal to μ no matter which samples are drawn.

$$2. (a) \quad E(\bar{Y}) = \frac{1}{n} E(Y_1 + \dots + Y_n) = \frac{E(Y_1) + \dots + E(Y_n)}{n} = \frac{\mu + \dots + \mu}{n} = \mu$$

$$(b) \quad \text{Var}(\bar{Y}) = \text{Var}\left(\frac{Y_1 + \dots + Y_n}{n}\right) = \frac{1}{n^2} \left[\text{Var}(Y_1) + \dots + \text{Var}(Y_n) + \sum_{i \neq j} \text{Cov}(Y_i, Y_j) \right] = \frac{\sigma^2}{n}$$

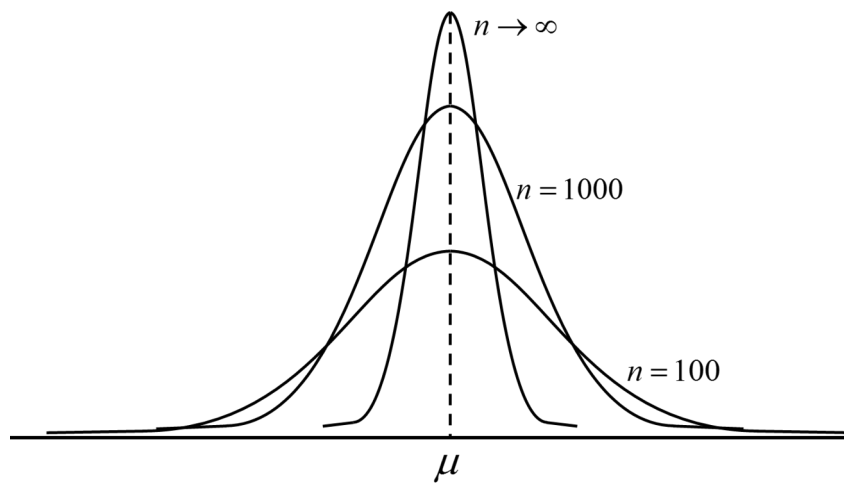
3. No. Failing to reject the null hypothesis means that the observed value of a test statistic is very likely to occur under $H_0: \mu = 15$. However, the same observed test statistic is also very likely to occur under slightly different null hypotheses, e.g., $H_0: \mu = 15.0001$. We cannot tell which one ($\mu = 15$ or $\mu = 15.0001$) is true based on such sample information.

4. To maximize the probability of correctly rejecting the null hypothesis when the alternative is true, the rejection region should be consisted of the values which are least probable under the null hypothesis but instead more probable under the alternative hypothesis. If the alternative of $H_1: \mu < 28$ is correct, the observed sample mean will be small. Thus we can predict that the smaller the observed sample mean is, the less likely the null is supported. Therefore, we need to set up the rejection region using the smallest values until the rejection region is equal to a given significance level.



5.

(a) Yes, for most estimators considered in econometric analysis.



However, if the variance of an unbiased estimator does not decrease to zero for an infinitely large sample, the estimator is not consistent. (This is a rare case, though.)

(b) No.