1

A

True.

Because the P-value of a hypothesis test is the probability of getting sample data at least as inconsistent with the null hypothesis (and supportive of the alternative hypothesis) as the sample data actually obtained. So small P-values provide evidence against the null hypothesis. In other words, the smaller the P-value is, the more likely that the sample value falls in the rejection region.

B

Uncertain.

Because we don't know the distribution of population and the definition of variability. The $\bar{X}$ is an unbiased estimator but without enough information we can't conclude it as efficient estimator.

C

False,

If $\bar{\theta} = \bar{X} + n$ is a estimator of X (n is the sample size), then

$MSE(\bar{\theta}) = Var(\bar{X} + n) + (E(\bar{X}) - X)^2 = \frac{Var^2(X)}{n} + n^2$

if n is large enough, then the $MSE(\bar{\theta})$ of large sample size will be higher than those of lower sample size.

D

False.

An estimator whose standard error is smaller than those of other potential estimators is said to be efficient but not consistent.

2

A

$$E(\hat{\theta}_1 - \hat{\theta}_2) = E(\hat{\theta}_1) - E(\hat{\theta}_2) = \theta_1 - \theta_2$$

B

$E(\hat{\theta}_1^2) = D(\hat{\theta}_1) + E^2(\hat{\theta}_1) = D(\hat{\theta}_1) + \theta_1^2$

Since $D(\hat{\theta}_1) \neq 0$, $E(\hat{\theta}_1^2) \neq \theta_1^2$

So $E(\hat{\theta}_1^2)$ is not an unbiased estimator of θ_1^2

C

Uncertain.

If the joint distribution of X and Y is:

Y \ X	1	0
1	$\frac{1}{4}$	$\frac{1}{4}$
0	$\frac{1}{4}$	$\frac{1}{4}$

$\hat{\theta}_1 = \bar{X}, \hat{\theta}_2 = \bar{Y}$.

Then $\hat{\theta}_1 = \bar{X}$ is the unbiased estimator of X, and $\hat{\theta}_2 = \bar{Y}$ is the unbiased estimator of Y. Under this condition, $\hat{\theta}_1 * \hat{\theta}_2$ is the unbiased estimator of $X * Y$.

$E(\hat{\theta}_1 * \hat{\theta}_2) = E(X * Y) = \frac{1}{4}$

If the joint distribution of X and Y is:

Y \ X	1	0
1	0	$\frac{1}{2}$
0	$\frac{1}{2}$	0

Then $\hat{\theta}_1 = \bar{X}$ is the unbiased estimator of X, and $\hat{\theta}_2 = \bar{Y}$ is the unbiased estimator of Y. Under this condition, $\hat{\theta}_1 * \hat{\theta}_2$ is not the unbiased estimator of $X * Y$.

$E(X * Y) = 0, E(\hat{\theta}_1 * \hat{\theta}_2) \neq 0$

3

A

Parameter: the statistical feature of the population's distribution that need to be estimated.

Estimators: An estimator is a statistic that estimates some feature about the population.

Estimates: Estimates is the value that calculated by applying the sample to the estimator.

B

Unbiased estimator:

In statistics, the bias (or bias function) of an estimator is the difference between this estimator's expected value and the true value of the parameter being estimated. An estimator with zero bias is called unbiased. In other words, the expected value of the estimator matches that of the parameter.

Consistent estimator:

In statistics, a consistent estimator is an estimator having the property that as the number of data points used increases indefinitely, the resulting sequence of estimates converges in probability to the parameter being estimated.This means

that the distributions of the estimates become more and more concentrated near the true value of the parameter being estimated

Efficient estimator:

An efficient estimator is an estimator that estimates the quantity of interest in some best possible manner. The notion of best possible relies upon the choice of a particular loss function the function which quantifies the relative degree of undesirability of estimation errors of different magnitudes. A more efficient estimator needs fewer observations than a less efficient one to achieve a given performance.

4

A

$$E(\hat{s}^2) = E(\frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n})$$

$$= E(\frac{\sum_{i=1}^{n} X_i^2}{n}) - E(\bar{X}^2)$$

$$= E(X^2) - E(\bar{X}^2)$$

for $E(\bar{X}) = E(X)$, $D(\bar{X}) = \frac{D(X)}{n}$

$E(\bar{X}^2) = D(\bar{X}) + E^2(\bar{X}) = \frac{D(X)}{n} + E^2(X)$

$E(\hat{s}^2) = E(X^2) - (\frac{D(X)}{n} + E^2(X)) = (1 - \frac{1}{n})D(X) \neq D(X)$

B

bias$= |E(\hat{s}^2) - D(X)| = \frac{1}{n}D(X)$

C

For:

$$\lim_{n\to\infty} bias = \lim_{n\to\infty} \frac{1}{n}D(X) = 0$$

So this estimator is asymptotically unbiased

5

A

B

6

$H_0 = 3$ black & 3 white $H_1 = 2$ black & 4 white

The probability of a Type 1:
$P(2\ white|H_0) = \frac{3}{6} \times \frac{2}{5} = \frac{1}{5}$

The probability of a Type 2:
$P(have\ black|H_1) = 2 \times \frac{2}{6} \times \frac{4}{5} + \frac{2 \times 1}{6 \times 5} = \frac{3}{5}$

7

A

True.

For a fixed sample size, the smaller we specify the significance level, α, the larger will be the probability of a Type 2 error, β, of not rejecting a false null hypothesis.

B

True.

The P-value of a hypothesis test equals the smallest significance level at which the null hypothesis can be rejected, that is, the smallest significance level for which the observed sample data results in rejection of H_0. In other words, the smaller (closer to 0) the P-value, the stronger is the evidence against the null hypothesis and, hence, in favor of the alternative hypothesis.

C

False.

If the P-value is larger than 0.01 and smaller than 0.05, then the null hypothesis is rejected against an alternative at the 5% level, but it would not be rejected against that alternative at the 1% level.

8

The nonrejection region of One-Mean z-Test I_z is:

$$\left(\mu_0 - \frac{z_{\alpha/2}}{\sigma/\sqrt{n}}, \mu_0 + \frac{z_{\alpha/2}}{\sigma/\sqrt{n}}\right)$$

The nonrejection region of corresponding t-Test I_t is:

$$\left(\mu_0 - \frac{t_{\alpha/2}}{s/\sqrt{n}}, \mu_0 + \frac{t_{\alpha/2}}{s/\sqrt{n}}\right)$$

since $t_{\alpha/2} > z_{\alpha/2}$

So I_z is the prime subset of I_t.

Given that the null hypothesis would have not been rejected had σ been known, the value of the estimator would be in I_z. So the value of estimator

would also be in I_t, which means it is impossible to reject the null hypothesis when is unknown under this condition.