Introduction to Econometrics
Sogang University

In Choi
Autumn, 2017

Home Assignment 1

Due: September 21, 2017

1. Consider the experiment of tossing two fair coins. Define the random variable X as
$$X = \frac{\textit{number of heads + 1}}{\textit{number of tails + 1}}.$$
   (a) Is the random variable X discrete or continuous? Explain.
   (b) Tabulate the probability distribution of the random variable X.
   (c) Calculate $E(X)$ and $Var(X)$.

2. The random variable X has the following probability distribution.

x	0	4	5	6
$p(x)$	.1	.3	.3	?

   (a) What is the probability $P(X = 6)$?
   (b) Calculate $E(X)$ and $Var(X)$.

3. Suppose that the random variable Z has a standard normal distribution. Calculate the following probabilities.
   (a) $P(Z > 1)$;
   (b) $P(Z < -1.5)$;
   (c) $P(\mid Z \mid > 1.7)$.

4. Suppose that the random variable Z has a standard normal distribution.
   (a) Find z_0 such that $P(\mid Z \mid < z_0) = .9$.
   (b) Find z_0 such that $P(Z > z_0) = .1$.
   (c) Find z_0 such that $P(Z < z_0) = .05$.

5. Suppose that $X \sim N(1, 9)$.
   (a) Find $P(X > 1)$.
   (b) Find $P(X > -2)$.

(c) Find $P(X < 4)$.

6. Suppose that $X_i \sim B(1, \frac{1}{3})$.

   (a) What is the sampling distribution of $\bar{X} = \frac{1}{2}\sum_{i=1}^{2} X_i$?

   (b) Show that $E(\bar{X}) = E(X_i)$.

7. A random sample of $n = 100$ observations is selected from a population having a binomial distribution $B(4, \frac{1}{2})$. What is the approximate value of the probability $P(\bar{X} < 2.1)$?

8. Let $X_i \sim N(3, 5)$ for all i. What is the sampling distribution of $\bar{X} = \frac{1}{5}\sum_{i=1}^{5} X_i$?

9. Let $X_i \sim N(3, 5)$ for all i. Calculate the probability $P(\frac{1}{5}\sum_{i=1}^{5} X_i < 4)$?

10. Let $X_i \sim N(3, 5)$ for all i.

    (a) What is the variance of $\bar{X} = \frac{1}{3}\sum_{i=1}^{3} X_i$?

    (b) What is the variance of $\bar{X} = \frac{1}{5}\sum_{i=1}^{5} X_i$?

    (c) How does the variance of the sample mean change as sample size increases? Infer from the answers for parts (a) and (b).

11. Discuss the validity of the following statements.

    (a) The central limit theorem for the sample mean works well even when the sample size is small.

    (b) The central limit theorem for the sample mean works only when the underlying distribution is normal.

    (c) The smaller the variance of the sample mean is, the more accurate it is.

    (d) When the underlying distribution is not normal, the sample mean has a normal distribution at any sample sizes.

12. Discuss the validity of the following statements.

    (a) A 90% confidence interval's length is greater than that of the corresponding 95% confidence interval.

    (b) To construct large sample confidence intervals for a population mean, a normality assumption is always required.

    (c) Confidence intervals for a population proportion are based on a normality assumption.

13. A random sample of 90 observations produced a mean $\bar{x} = 25.9$ and a standard deviation $s = 2.7$.

    (a) Find a 95% confidence interval for the population mean μ without assuming normality.

    (b) How does the answer for part (a) change if we assume that the sample is from a normal distribution.

    (c) Find a 99% confidence interval for the population mean μ without assuming normality.

    (d) Compare the widths of confidence intervals from parts (a) and (c); and interpret the result.

14. A company is interested in estimating μ, the mean number of days of sick leave taken by all its employees. The firm's statistician selects at random 10 personnel files and notes the number of sick days taken by each employee. The following sample statistics are computed: $\bar{x} = 12.2$ days, $s = 10$ days.

    (a) Estimate μ by using a 90% confidence interval without making a normality assumption.

    (b) Estimate μ by using a 90% confidence interval by making a normality assumption.

    (c) Which confidence interval do you think is more accurate?

15. Discuss the validity of the following statements.

    (a) We may reduce both Type I and Type II errors at a given sample size.

    (b) Larger p-values imply stronger support for the null hypothesis.

    (c) Type I error is the probability of rejecting the null hypothesis when it is true.

    (d) Type I error cannot be controlled in hypothesis testing.

16. A random sample of 25 observations produced the following summary statistics: $\bar{x} = 0.323$ and $s^2 = 0.034$.

    (a) Test the null hypothesis $H_0 : \mu = 0.36$ against the alternative $H_1 : \mu < 0.36$ at the 10% level.

    (b) Test the null hypothesis $H_0 : \mu = 0.36$ against the alternative $H_1 : \mu \neq 0.36$ at the 10% level.

17. In a test of the hypothesis $H_0 : \mu = 100$ against $H_1 : \mu \neq 100$, the sample data yielded the test statistic $z = 2.17$. Find the p-value for the test.

18. A population we are interested in has a normal distribution with mean μ and unknown variance. A t-test is conducted for the null hypothesis $H_0 : \mu = 10$ against $H_1 : \mu > 10$ for a random sample of 17 observations. The computed value of the t-test is 1.174. Do we reject the null at the 5% significance level? Use the small-sample approach.