Table 1: Table for Exercise 1.

VICTIMS	Frequency	Percent
0	1244	90.8
1	81	5.9
2	27	2.0
3	11	0.8
4	4	0.3
5	2	0.1
6	1	0.1

N	Mean	Std Dev
1370	0.146	0.546

PRACTICE QUESTIONS FOR EXAM 1

1. The 1990 General Social Survey asked respondents, "During the past 12 months, how many people have you known personally that were victims of homicide." Table 1 shows a computer printout from analyzing responses for 1370 subjects.
   a. Does the Empirical Rule apply to this distribution. Why or why not?
   b. Based on the mean and standard deviation, explain what you would surmise about the shape of the distribution.

2. Find the z-score for the upper quartile of a normal distribution.

3. According to *Current Population Reports*, self-employed individuals in the United States in 1990 worked an average of 44.6 hours per week, with a standard deviation of 14.5. Assuming this variable is approximately normally distributed, find the proportion who averaged more than 40 hours per week.

4. The 1991 General Social Survey asked, "During the last year, did anyone take something from you by using force – such as a stickup, mugging, or threat?" Of 987 subjects, 17 answered yes and 970 answered no.

   a. Construct a 95% confidence interval for the population proportion who would answer yes.

   b. Interpret the interval in (a).

5. A significance test is conducted about the value of a population mean, to see whether it differs from 100. The sample of 100 observations have a mean of 97 and a standard deviation of 30.

   a. Define notation, and state the hypotheses.

   b. Find the test statistic.

   c. Find the P-value, and interpret.

   d. For the α-level of .05, what would be your conclusion about the null hypothesis?

For the following multiple-choice items, select the correct response(s).

6. (5 pts.) The standard error of a statistic describes

   a. The standard deviation of the sampling distribution of that statistic.

   b. The standard deviation of the sample measurements.

   c. How close that statistic is likely to fall to the parameter that it estimates.

   d. The variability in the values of the statistic for repeated random samples of size n.

   e. The error that occurs due to nonresponse and measurement errors (such as lying to the interviewer).

7. (4 pts.) The Central Limit Theorem implies that

   a. All variables have approximately bell-shaped sample distributions if a random sample contains at least 30 observations.

   b. Population distributions are normal whenever the population size is large.

   c. For large random samples, the sampling distribution of $\overline{y}$ is approximately normal, regardless of the shape of the population distribution.

   d. The sampling distribution looks more like the population distribution as the sample size increases.

   e. All of the above.

8. (5 pts.) Based on responses of 1467 subjects in General Social Surveys in the mid-1980s, a 95% confidence interval for the mean number of close friends equals (6.8, 8.0). Which of the following interpretations is (are) correct?

   a. We can be 95% confident that $\overline{Y}$ is between 6.8 and 8.0.

   b. We can be 95% confident that μ is between 6.8 and 8.0.

   c. Ninety-five percent of the values of Y = number of close friends (for this sample) are between 6.8 and 8.0.

d. If random samples of size 1467 were repeatedly selected, then 95% of the time $\overline{Y}$ would be between 6.8 and 8.0.

e. If random samples of size 1467 were repeatedly selected, then in the long run 95% of the confidence intervals formed would contain the true value of μ.

Indicate true (T) or false (F):

9. Outliers have a greater impact on the mean than the median. ______

10. We want to determine the sample size needed to obtain a good estimate of a population proportion π. To be "conservative" in determining the answer, in terms of having a sample size that is safe even if somewhat larger than actually needed, we can assume that $\pi = 0$ or $\pi = 1$. ______

Solutions

1. a. No, distribution not even approximately bell-shaped. b. Highly skewed to the right.

2. z = 0.67

3. About 0.62

4.a. 0.017 $\pm$ 0.008, or (0.009, 0.025). b. We can be 95% confident that the population proportion who would answer yes is between 0.009 and 0.025.

5.a. H_0: $\mu = 100$, H_a: $\mu \neq 100$. b. $t = -1.0$. c. $P = 0.32$. If H_0 is true, the probability would be 0.32 of getting a sample mean at least as far from 100 as the observed sample mean (or, a t statistic at least 1.0 in absolute value). d. Do not reject H_0. It is plausible that $\mu = 100$.

6. a, c, d

7. c

8. b, e

9. T

10. F

Formulas

$$\overline{y} = \frac{\Sigma y_i}{n} \qquad s^2 = \frac{\Sigma(y-\overline{y})^2}{n-1}$$

$$\sigma_{\overline{y}} = \frac{\sigma}{\sqrt{n}} \qquad z = \frac{y-\mu}{\sigma}$$

$$\overline{y} \pm t(se), \quad se = s/\sqrt{n}$$

$$t = \frac{\overline{y}-\mu_0}{se}$$

$$\sigma_{\hat{\pi}} = \sqrt{\frac{\pi(1-\pi)}{n}}$$

$$\hat{\pi} \pm z(se), \quad se = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

$$z = \frac{\hat{\pi}-\pi_0}{se_0} \qquad se_0 = \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$$

$$n = \left(\frac{z}{M}\right)^2 \pi(1-\pi)$$

$$n = \left(\frac{z}{M}\right)^2 \sigma^2$$