Biostatistics 140.621

More Extra Problems

1. Circle the correct response for EACH statement:

T F A parameter describes the distribution of a population of observations.

T F A dichotomous event has two equally likely outcomes.

T F Mutually exclusive events are independent events.

T F The binomial distribution always assumes independent events.

T F The Central Limit Theorem assures that, with small enough samples, the sampling distribution of a sample statistic is approximately normal.

T F When two events A and B are independent, the probability of both events occurring together is zero.

T F A probability distribution is a relative frequency distribution.

T F A statistic is an estimate.

Problems 2a through 2d concern the following: In a certain Village A, it is hypothesized that the probability of a child being malnourished is 0.04. There are 1,000 children in Village A.

2a. Which probability distribution would best describe the number of malnourished children in the population? (Check only one response.)

( ) a. Binomial with p=0.4 , n=1,000
( ) b. Poisson with λ = 0.04
( ) c. Normal with mean = 4 and standard deviation = 0.006
( ) d. Binomial with p =0.04, n = 1,000
( ) e. Chi-square with 999 degrees of freedom

2b. What assumptions are made with your response to problem 2a? (Check ALL responses that apply.)

( ) a. Independence of occurrences of malnutrition among children.
( ) b. Same probability of malnutrition for each child.
( ) c. Different probabilities of malnutrition for each child.
( ) d. Malnutrition is a continuous measurement.
( ) e. Malnutrition is a dichotomous event.

2c. Assume that the probability of being malnourished in a neighboring Village B is the same as that of village A. Suppose it was found that 3 children out of a random sample of 50 children in Village B are malnourished. How likely is it to observe such a large proportion? (Check ALL responses that apply.)

( ) a. $P(Z > 0.72)$
( ) b. $P(Z > 3.23)$
( ) c. $P(\hat{p} > 0.06)$
( ) d. $P(\hat{p} > 0.04)$
( ) e. $P(Z > 0.06)$

2d. Based on the data and your response to problem 2c, what do you conclude regarding the probability of malnutrition in Village B? (Check only one response.)

( ) a. It appears to be higher than that of Village A.
( ) b. It appears to be lower than that of Village A.
( ) c. It appears to be the same as that of Village A.
( ) d. No conclusion can be made.

Problems 3a through 3c concern the following: Suppose that the distribution of serum Vitamin E is approximately normal with a mean of 860 ug/dl and standard deviation of 340 ug/dl.

3a. If the "normal range" is considered to be within (+/-) 1.5 standard deviations of the mean, this range is between: (Check only one response.)

( ) a. 520 - 1200 ug/dl
( ) b. 180 - 1540 ug/dl
( ) c. 350 - 1370 ug/dl
( ) d. 690 - 1030 ug/dl
( ) e. 340 - 860 ug/dl

3b. What is the probability of an individual with a Vitamin E level falling outside of the normal range? (Check only one response.)

( ) a. 0.13
( ) b. 0.05
( ) c. 0.023
( ) d. 0.15
( ) e. 0.32

3c. What is the probability of an individual with a Vitamin E level more than two standard deviations above the mean value? (Check only one response.)

( ) a. 0.13
( ) b. 0.05
( ) c. 0.023
( ) d. 0.15
( ) e. 0.32

4. Circle either True or False for EACH statement:

T F A parameter provides an estimate of a statistic.

T F Both the Poisson and normal distributions may be used to approximate the binomial distribution.

T F The binomial distribution assumes the same probability of "success" across observations (trials).

T F The Central Limit Theorem assures that, with large enough samples, the population distribution is approximately normal.

T F When two events A and B are mutually exclusive, the probability of both events occurring together is one.

T F A statistic is calculated from a population.

T F A characteristic of the normal distribution is its symmetry about its median.

T F The standard error of the mean is the same as the sample standard deviation.

T F For a continuous probability distribution, $P(X < x) = P(X \leq x)$.

T F For a discrete probability distribution, $P(X < x) = P(X \leq x)$.

Problems 5a through 5d concern the following: Suppose that albumin follows a normal distribution in a healthy population with $\mu = 3.75$ mg/100 ml and $\sigma = 0.50$ mg/100 ml. The normal range of values will be defined as $\mu \pm 1.96\ \sigma$, so that values outside these limits are classified as "abnormal".

5a. Abnormal values are: (Check only one response.)

( ) a. < 3.25 and > 4.25
( ) b. < 2.77 and > 4.73
( ) c. < -1.96 and > 1.96
( ) d. < -0.98 and > 0.98

5b. What is the probability of values of 2.5 mg/100 ml or less? (Check only one response.)

( ) a. 0.9938
( ) b. 0.9970
( ) c. 0.0048
( ) d. 0.0062
( ) e. 0

5c. Suppose the mean albumin value in a sample of 25 patients with advanced chronic liver disease is 2.5 mg/100 ml. What is the probability of this value or something smaller? (Check only one response.)

( ) a. 0.9938
( ) b. 0.9970
( ) c. 0.0030
( ) d. 0.0062
( ) e. 0

5d. Based on the answer to Problem 5c, one would conclude: (Check only one response.)

( ) a. It is unlikely that this sample came from a population with a mean albumin value of 3.75.

( ) b. It is likely that this sample came from a population with a mean albumin value of 3.75.

( ) c. It is likely that this sample came from a population with a mean albumin value of 2.5.

( ) d. It is unlikely that this sample came from a population with a mean albumin value of 2.5.

Problems 6a through 6c concern the following: The annual mortality rate for children ages 1-4 years in Bangladesh in 17 per 1,000.

6a. Given this information, which probability distribution would best describe the number of deaths in this age group in Bangladesh? (Check only one response.)

( ) a. Poisson
( ) b. Normal
( ) c. Binomial
( ) d. Chi-square

6b. What is the probability that none will die in a random group of 50 1-4 year-olds? (Check only one response.)

( ) a. 0.998
( ) b. 0.607
( ) c. 0.183
( ) d. 0.427

6c. The answer to problem 4b can be solved using either: (Check only one response.)

( ) a. The Poisson distribution with λ=0.50 or the binomial distribution with p=0.017

( ) b. The Poisson distribution with λ=0.85 or the binomial distribution with p=0.017

( ) c. The Poisson distribution with λ= 0.017 or the normal approximation to the binomial distribution.

( ) d. The Poisson distribution with λ=0.017 or a binomial distribution with p=0.34.

Problems 7a through 7d concern the following: Research has concluded that individuals experience a common cold approximately two times per year. Assume that the time between colds is normally distributed with a mean of 160 days and a standard deviation of 40 days.

7a. What is the probability of going 200 or more days between colds? (Check only one response.)

( ) a. 0.682
( ) b. 0.138
( ) c. 0.841
( ) d. 0.159
( ) e. 0.318

7b. Approximately 95% of this population will experience colds between: (Check only one response.)

( ) a. 0 and 320 days
( ) b. 120 and 200 days
( ) c. 40 and 280 days
( ) d. 80 and 240 days

7c. Suppose that the mean time between colds in a random sample of 100 individuals is 150 days. What is the probability of this value or something smaller? (Check only one response.)

( ) a. 0.006
( ) b. 0.401
( ) c. 0.994
( ) d. 0.599
( ) e. 0

7d. Based on the answer to Problem 7c, one would conclude: (Check only one response.)

( ) a. It is unlikely that this sample came from a population with a mean of 150 days between colds

( ) b. It is likely that this sample came from a population with a mean of 160 days between colds

( ) c. It is likely that this sample came from a population with a mean of 150 days between colds

( ) d. It is unlikely that this sample came from a population with a mean of 160 days between colds.

Suppose that the distribution of total scores for a class of biostatistics students is approximately normally distributed with a mean of 85 points and a standard deviation of 9 points. 8

8a. The probability of an individual scoring above 93 is: (Check only one response.)

( ) a. 0.889
( ) b. 0.110
( ) c. 0.187
( ) d. 0.813
( ) e. 0.320

8b. What is the probability of an individual scoring lower than one standard deviation below the mean? (Check only one response.)

( ) a. 0.889
( ) b. 0.841
( ) c. 0.159
( ) d. 0.111
( ) e. 0.320

8c. What is the probability of an individual scoring more than 1.5 standard deviations above the mean value? (Check only one response.)

( ) a. 0.067
( ) b. 0.933
( ) c. 1.000
( ) d. 0.477
( ) e. 0.889

8d. Suppose that the mean score in a random sample of 36 students is 87. What is the probability of this value or something smaller? (Check only one response.)

( ) a. 0.889
( ) b. 0.092
( ) c. 0.413
( ) d. 0.587
( ) e. 0.908

8e. Suppose a random sample of 36 students is chosen. What is the probability that at least 25 percent score higher than 93? (Check only one response.)

( ) a. 0.918
( ) b. 0. 79
( ) c. 0.821
( ) d. 0.166
( ) e. 0.889

Problems 9a through 9c concern the following: It is observed that, on average, five smokers gather to smoke on the front steps of the School of Public Health during one a one hour period. A smoking-awareness campaign is initiated.

9a. Suppose that after the campaign had ended, 8 smokers were observed on the front steps during a three hour period: The probability of this observation or something smaller is: (Check only one response.)

( ) a. 0.037
( ) b. 0.118
( ) c. 0.882
( ) d. 0.070
( ) e. 1.000

9b. What would you conclude from your response in problem 6a? Was the smoking- awareness campaign effective? (Check only one response.)

( ) a. It is unlikely that this campaign was effective; the observed number of smokers was greater than what would have been expected before the campaign took place.
( ) b. It is likely that this campaign was effective; the observed number of smokers was less than what would have been expected before the campaign took place.
( ) c. It is unlikely that this campaign was effective; the observed number of smokers was the same as what would have been expected before the campaign took place
( ) d. No conclusion can be made.

9c. What probability distribution was used to describe the number of smokers in these problems? (Check only one response.)

( ) a. Binomial with $p = 0.001$
( ) b. Poisson with $\lambda = 15$
( ) c. Normal with mean = 0.001 and standard deviation = 0.003
( ) d. Poisson with $\lambda = 5$

Problems 10a through 10e concern the following: Suppose the mean duration of gestation (pregnancy) for a population of healthy women is 280 days with a standard deviation of 10 days.

10a. What proportion of women are "overdue" (longer than the mean duration) by more than one week?

( ) a. 0.7580
( ) b. 0.5398
( ) c. 0.2420
( ) d. 0.4602

10b. What proportion of women are "overdue" (longer than the mean duration) by more than two weeks?

( ) a. 0.0808
( ) b. 0.5438
( ) c. 0.9192
( ) d. 0.4562

10c. Suppose a sample of 4 women with a particular condition was selected from this population and their respective lengths of gestation were:

240 days
250 days
270 days
280 days

What is the sample mean? (Check only one response.)

( ) a. Shorter than the mean duration for the population
( ) b. Longer than the mean duration for the population
( ) c. Same as the mean duration for the population
( ) d. Cannot determine from this population

10d. What is the probability that the mean gestation is less than that observed for this sample? (Check only one response.)

( ) a. $P(\bar{X} > 260)$

( ) b. $P(Z < -2)$

( ) c. $P(\bar{X} > 20)$

( ) d. $P(Z < -4)$

10e. Is there evidence of a shorter gestation for this group? (Check only one response.)

( ) a. Yes, because the sample deviation is smaller
( ) b. Yes, the value of the sample mean is unlikely when the population mean is 280
( ) c. No, because the value of this sample mean is likely
( ) d. No, because a sample of size 4 has low probability

Problems 11a and 11b concern the following: Suppose that the mean number of serious accidents in a large workplace over a ten year period is 2.

11a. The number of serious accidents in a one-year period can be described by: (Check only one response.)

( ) a. Binomial distribution with p = 0.2
( ) b. Normal distribution with mean = 2 and standard deviation = 10
( ) c. Poisson distribution with λ = 0.2
( ) d. Chi-square distribution with degrees of freedom = 1

11b. The probability of more than one serious accident in a one-year period is: (Check only one response.)

( ) a. 0.982
( ) b. 0.163
( ) c. 0.406
( ) d. 0.018
( ) e. 0.594

12. Fill in the appropriate response for each blank:

a. Probabilities can take on values greater than or equal to _________ and less than or equal to ________.
b. When graphed, a probability distribution for discrete data takes the form of a _______________.
c. The probability of a single point under a continuous distribution equals _______.
d. The Poisson distribution approximates the __________ distribution when n is large.
e. The area under the curve for a continuous distribution integrates to __________.

NAME ______________________________

Problems 13a through 13f concern the following: Suppose that in a certain maternity hospital, the length of stay for a mother follows a normal distribution with a mean of 3 days and a standard deviation of one day.

13a. What is the probability of a mother going home in less than one day? (Check only one response.)

( ) a. $P(X < 1)$
( ) b. $P(Z < -2)$
( ) c. $P(Z < 2)$
( ) d. a is the same as b
( ) e. a is the same as c

13b. The above probability in problem 3a is equal to: (Check only one response.)

( ) a. 0.0228
( ) b. 0.0183
( ) c. 0.9772
( ) d. 0.9817

13c. What is the probability of a mother going home between 1 and 2 days? (Check only one response.)

( ) a. $P(2 < X < 1)$
( ) b. $P(1 < Z < 2)$
( ) c. $P(-2 < Z < -1)$
( ) d. a and b

13d. The above probability in problem 3c is equal to: (Check only one response.)

( ) a. 0.1756
( ) b. 0.9772
( ) c. 0.8185
( ) d. 0.1359

13e. Suppose a random sample of 10 mothers was taken from the total population of obstetric admissions. What is the probability that the mean length of stay is greater than two days? (Check only one response.)

( ) a. $P(\bar{X} > 10)$

( ) b. $P(\bar{X} > 2/\sqrt{10})$
( ) c. $P(Z < -3.16)$
( ) d. $P(Z > -3.16)$

13f. The above probability in problem 13e is equal to: (Check only one response.)

( ) a. 0.0228
( ) b. 0
( ) c. 0.0008
( ) d. 0.9992

NAME________________________

Problems 14a through 14d concern the following:

In a certain village, the probability of a child having xerophthalmia (Vitamin A deficiency) is 0.04. There are 50 children in the village.

14a. Which probability distribution would best describe the number of children with xerophthalmia in this population? (Check only one response.)

( ) a. Binomial; $p = 0.04$, $n = 10$
( ) b. Poisson; $\lambda = 4$
( ) c. Poisson; $\lambda = 2$
( ) d. Binomial; $p = 0.10$, $n = 50$

14b. What is the probability that at least 10% of the children have xerophthalmia? (Check only one response.)

( ) a. 0.983
( ) b. 0.053
( ) c. 0.215
( ) d. 0.017

14c. What is the probability that none have xerophthalmia? (Check only one response.)

( ) a. 0.271
( ) b. 0.135
( ) c. 0.018
( ) d. 0.729

14d. If 25 children were observed to have xerophthalmia, what would you conclude? (Check only one response.)

( ) a. The probability of such an observation is zero
( ) b. The probability of observing this is approximately zero if the true probability of xerophthalmia is 0.04
( ) c. The probability of such an observation is one
( ) d. The probability of observing this is greater than the probability of observing 20 children with xerophthalmia

15. In a study of the quality of life for cancer patients, the following variables were measured. Please match each variable with the probability distribution that may describe it (MULTIPLE CHOICES ALLOWED but indicate your best choice)

a. Poisson b. Normal c. Binomial

____ Assistance in daily living (1=yes, 0=no)

____ Age (in years)

____ Number of attempted suicides

____ Time since the diagnosis of cancer (in days)

____ Gender (1=male, 0=female)

____ Weight (in kg)

____ Other adults in household (1=yes, 0=no)

____ Number of days without nausea during treatment

16. Circle a response of True (T) or False (F) for EACH statement:

T F The significance level of a hypothesis test is the probability of rejecting the null hypothesis when it is true.

T F A statistical decision is always correct.

T F The t-test may be used for comparing means or proportions between two groups.

T F Power is the probability of not detecting a difference when a difference really exists.

T F A p-value is calculated under the assumption that the null hypothesis is false.

T F A sample statistic provides an estimate of a population parameter.

T F Construction of a confidence interval does not require the statement of a null hypothesis.

17. A randomized trial was conducted in Belgium to compare drug treatment based on conventional blood pressure measurements (CBP) with drug treatment based on ambulatory blood pressure measurements (ABP) in the management of hypertension (JAMA 1997;278:1065). The following table shows the baseline characteristics of patients in both treatment groups:

Table 1.—Baseline Characteristics of Patients Randomized to Antihypertensive Drug Treatment Based on Conventional Blood Pressure (CBP) or Ambulatory Blood Pressure (ABP) Measurements

Characteristics	CBP Group (n=206)	ABP Group (n=213)	*P*
Age, mean (SD), y	51.3 (11.9)	53.8 (10.8)	.03
Body mass index, mean (SD), kg/m²	28.5 (4.8)	28.2 (4.4)	.39
Women, No. (%)	102 (49.5)	124 (58.2)	.07
Receiving oral contraceptives, No. (%)*	14 (13.7)	10 (8.1)	.17
Receiving hormonal substitution, No. (%)*	19 (18.6)	19 (15.3)	.51
Previous antihypertensive treatment, No. (%)†	134 (65.0)	139 (65.3)	.95
Diuretics, No. (%)*	47 (35.1)	59 (42.4)	.26
β-Blockers, No. (%)*	65 (48.5)	80 (57.6)	.17
Calcium channel blockers, No. (%)*	45 (33.6)	38 (27.3)	.32
Angiotensin-converting enzyme inhibitors, No. (%)*	50 (37.3)	48 (34.5)	.72
Multiple-drug treatment, No. (%)*	62 (46.3)	65 (46.8)	.97
Smokers, No. (%)	42 (20.5)	35 (16.4)	.29
Alcohol use, No. (%)	115 (55.8)	102 (47.9)	.10
Serum creatinine, mean (SD), μmol/L‡	85.75 (15.91)	88.4 (16.80)	.25
Serum total cholesterol, mean (SD), mmol/L‡	6.00 (1.03)	6.10 (1.19)	.32

*Percentages and values of *P* computed considering only women receiving antihypertensive drug treatment before their enrollment.
†Defined as antihypertensive drug treatment within 6 months before the screening visit.
‡Divide creatinine by 88.4 and cholesterol by 0.02586 to convert milligrams per deciliter.

The STATA log shows the following for a test of differences in mean age by treatment:

```
12. sdtesti 206 51.3 11.9 213 53.8 10.8

Two-sample test of variance                      x: Number of obs =      206
                                                 y: Number of obs =      213

------------------------------------------------------------------------------
Variable |      Mean    Std. Err.        t     P>|t|      [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |      51.3    .8291123    61.8734   0.0000      49.66532    52.9346
       y |      53.8    .7400038    72.7023   0.0000      52.34129    55.258
------------------------------------------------------------------------------
combined |  52.57088    .5546837    94.7763   0.0000      51.48057    53.66
------------------------------------------------------------------------------

                        Ho: sd(x) = sd(y)
              F Observed   = F   = F(212,205) =     1.214
              F Lower tail = F_L = F(212,205) =     0.824
              F Upper tail = F_U = F(212,205) =     1.214

   Ha: s1 < s2               Ha: s1 ~= s2                Ha: s1 > s2
  P < F = 0.9186       P < F_L + P > F_U  = 0.1622      P > F = 0.0814
```

y: Number of obs = 213

13. ttesti 206 51.3 11.9 213 53.8 10.8

Variable	Mean	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	51.3	.8291123	61.8734	0.0000	49.66532	52.93468
y	53.8	.7400038	72.7023	0.0000	52.34129	55.258
diff	-2.5	1.109522	-2.25322	0.0248	-4.680954	-.3190463

Degrees of freedom: 417

Ho: mean(x) - mean(y) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = -2.2532	t = -2.2532	t = -2.2532
P < t = 0.0124	P > |t| = 0.0248	P > t = 0.9876

x: Number of obs = 206
y: Number of obs = 213

14. ttesti 206 51.3 11.9 213 53.8 10.8, unequal

Variable	Mean	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	51.3	.8291123	61.8734	0.0000	49.66532	52.93468
y	53.8	.7400038	72.7023	0.0000	52.34129	55.25871
diff	-2.5	1.11132	-2.24958	0.0250	-4.684595	-.3154046

Satterthwaite's degrees of freedom: 410.06796

Ho: mean(x) - mean(y) = diff = 0

Ha: diff < 0	Ha: diff ~= 0	Ha: diff > 0
t = -2.2496	t = -2.2496	t = -2.2496
P < t = 0.0125	P > |t| = 0.0250	P > t = 0.9875

17 a. **What does one conclude regarding the $H_0: \sigma_1^2 = \sigma_2^2$ (equal population variances)?: (Check only one response.)**

- ( ) a. **Reject the null hypothesis and pool sample variances**
- ( ) b. **Reject the null hypothesis and use individual sample variances**
- ( ) c. **Fail to reject the null hypothesis and pool sample variances**
- ( ) d. **Fail to reject the null hypothesis and use individual sample variances**

17 b. **The value of the test statistic of interest in testing $H_0: \mu_1 - \mu_2 = 0$ is: (Check only one response.)**

- ( ) a. **1.214**
- ( ) b. **-2.2532**
- ( ) c. **1.645**
- ( ) d. **1.96**
- ( ) e. **-2.496**

17c. The p-value corresponding to the test statistic is: (Check the most appropriate response.)

( ) a. 0.05
( ) b. 0.1622
( ) c. 0.0248
( ) d. 0.0250
( ) e. 0.0124

17d. The 95% confidence interval for the difference in mean age between treatment groups is calculated as: (Check only one response.)

( ) a. (-4.68,-0.32)
( ) b. (-5.36,0.36)
( ) c. (49.66,52.93)
( ) d. (52.34,55.26)
( ) e. (-2.61,-2.39)

17e. The 99% confidence interval for the difference in mean age between treatment groups may be calculated as: (Check only one response.)

( ) a. (-4.68,-0.32)
( ) b. (-5.36,0.36)
( ) c. (49.66,52.93)
( ) d. (52.34,55.26)
( ) e. (-2.61,-2.39)

17f. For this problem, using a 0.01 significance level, is there a difference in mean age between the two treatment groups? (Check only one response.)

( ) a. Group CBP has significantly lower mean age than Group ABP.
( ) b. Group CBP has significantly higher mean age than Group ABP.
( ) c. Mean ages do not differ significantly between Groups CBP and ABP.
( ) d. Mean ages are different between Group CBP and Group ABP.
( ) e. It is not possible to compare mean ages between the treatment groups.

18. In a randomized trial, 100 children received a new toothpaste A and 100 children received an available standard toothpaste B. The number of DMFs (decayed, missing and filled teeth) for each child was obtained after 3 years:

Group	Mean	SD
A	10.2	7.5
B	12.2	8.3

18a. An appropriate null hypothesis is: (Check only one response.)

( ) a. Ho: $\mu_1 = \mu_2$
( ) b. Ho: $\mu_1 \geq \mu_2$
( ) c. Ho: $\mu_1 \leq \mu_2$
( ) d. Ho: $\mu_d = 0$
( ) e. None of the above

18b. What is the rationale for your choice in 6a?

____

____

The STATA log shows the results of significance testing:

```
sdtesti 100 10.2 7.5 100 12.2 8.3

 Variable |       Obs          Mean    Std. Dev.
 ---------+-------------------------------------
        x |       100          10.2         7.5
        y |       100          12.2         8.3
 ---------+-------------------------------------
 combined |       200             .     7.91012

         Ho:       sd(x) = sd(y)    (two-sided test)
             F(99,99) = 1.22
             2*(Pr > F) = 0.3149

1. ttesti 100 10.2 7.5 100 12.2 8.3

 Variable |       Obs          Mean    Std. Dev.
 ---------+-------------------------------------
        x |       100          10.2         7.5
        y |       100          12.2         8.3
 ---------+-------------------------------------
 combined |       200          11.2    7.953653

         Ho:  mean(x) = mean(y)   (assuming equal variances)
                    t = -1.79 with 198 d.f.
             Pr > |t| = 0.0753
```

18c . For a significance level of 0.05 for this problem, the rejection region for a two-sided test is: (Check only one response.)

( ) a. $z < -1.96$ and $z > 1.96$
( ) b. $t < -1.984$ and $t > 1.984$
( ) c. $t < -1.660$ and $t > 1.660$
( ) d. $z < -1.645$ and $z > 1.645$

18d . Using a two-sided test, what could one conclude regarding the effect of toothpaste A in reducing DMFs? (Check only one response).

( ) a. Toothpaste A is significantly better than Toothpaste B.
( ) b. Toothpaste B is significantly better than Toothpaste A.
( ) c. Toothpaste A is not significantly different from Toothpaste B.
( ) d. No conclusion can be made.

18e For a one-sided test with a significance level of 0.05 for this problem: (Check ALL responses that apply.)

( ) a. The sample statistic is different from that of the two-sided test.
( ) b. The test statistic is different from that of the two-sided test.
( ) c. The rejection region is different from that of the two-sided test.
( ) d. The conclusion is different from that of the two-sided test.

18f Using a one-sided test, what could one conclude regarding the effect of toothpaste A in reducing DMFs? (Check only one response).

( ) a. Toothpaste A is significantly better than Toothpaste B.
( ) b. Toothpaste B is significantly better than Toothpaste A.
( ) c. Toothpaste A is not significantly different from Toothpaste B.
( ) d. No conclusion can be made.

19. Circle a response of True (T) or False (F) for EACH statement:

T F A one-sided test is always more conservative than a two-sided test.

T F A nonsignificant finding may be the result of small sample size.

T F A significant finding may be the result of large sample size.

T F There is 100% certainty that the true population parameter is contained within a 95% confidence interval.

ANSWERS:

1. T F F T F F T T

2a. d 2b. a, b, e 2c. a, c $H_0: p_B = p_A = 0.04 = p \quad q = 0.96$

$P(\hat{p} > 0.06) = P\left(z > \frac{0.06 - 0.04}{\sqrt{\frac{.04(.96)}{50}}}\right)$

2d. c. $= P(z > 0.72) = 0.236$ cannot reject H_0

3a. c 3b. a 3c. c $P(X > 860 + 2(340)) = P(z > 2)$

$\mu = 860 \quad \sigma = 340$

4. F T T F F F T F T F

5a. b. 5b. d $P(X < 2.5) = P\left(z < \frac{2.5 - 3.75}{0.5}\right) = P(z < -2.5)$

5c. e $P(\bar{X} < 2.5) = P\left(z < \frac{2.5 - 3.75}{0.5/\sqrt{25}}\right) = P(z < -12.5)$

5d. a.

6a. a since unknown population size, c if n is known

6b. d. 6c. b

7.a. d $P(X > 200) = P\left(z > \frac{200 - 160}{40}\right) = P(z > 1)$

7.b. d. 7c. a. $P(\bar{X} < 150) = P\left(z < \frac{150 - 160}{40/\sqrt{100}}\right) =$ 7d. d.

8.a. c 8b. c 8c. a 8.d. e $P(\bar{X} \le 87) =$

where $\mu = 85 \quad \sigma = 9$ $P\left(z \le \frac{87 - 85}{9/\sqrt{36}}\right) = P(z < 1.33)$

8.e. d $p = P(\text{scoring} > 93) = 0.187$

$\hat{p} = 0.25 \quad P(\hat{p} > 0.25) = P\left(z > \frac{0.25 - 0.187}{\sqrt{\frac{0.187(.813)}{36}}}\right) = P(z > 0.97)$

9.a. a 9b. b 9c. b. $\lambda = \mu \cdot t = 5/1hr \cdot 3hr$

$P(X \le 8)$ where $\lambda = 15$ $= 15$

10.a. c. $P(X > 287)$ $\mu = 280, \sigma = 10 \rightarrow P\left(z > \frac{287 - 280}{10}\right) = P(z > 0.7)$

10.b. a 10c. a. 10d. d. $P(\bar{X} < 260) = P\left(z < \frac{260 - 280}{10/\sqrt{4}}\right) = P(z < -4)$

10e. b. calculate from sample of 4

11.a. c. 11b. d.

12.a. 0, 1 12b. histogram 12c. 0 12d. binomial 12e. 1

13a. d $P(X < 1) = P\left(z < \frac{1 - 3}{1}\right) = P(z < -2)$

13b. a 13c. c. 13d. d. 13e. d $P(\bar{X} > 2) = P\left(z > \frac{2 - 3}{1/\sqrt{10}}\right) = P(z > -3.16)$

13f. d.

14.a. c $n = 50, p = 0.04 \rightarrow \lambda = np = 2$ 14b. b. $10\%(50) = 5$

14c. b 14d. b. $P(X = 25) \approx 0$ $P(X \ge 5) = 1 - P(X \le 4)$

15. c, b, a, b, c, b, c, a

16. T F F F F T T

17a. c. 17b. b. 17c. c 17d. a 17e. b 17f. c.

18.a. a 18b. a difference could be expected in either direction

18.c. a. 18d. c 18e. c&d 18f. a

19. F T T F

Binomial Formula

$$P(X=x) = \binom{n}{x} p^x q^{n-x}$$

Poisson Formula

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Normal Transformation (Standardization)

$$z = \frac{x-\mu}{\sigma}$$

Tables will be provided for z, t

Sampling Distributions

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

$$z = \frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}$$