STATISTICS 641 - FINAL EXAMINATION

Student's Name ____

Student's Email Address ____

INSTRUCTIONS FOR STUDENTS:

(1) The exam consists of 10 pages including this cover page, and 22 pages of Tables of distributions and percentiles.

(2) You have exactly **2 hours** to complete the exam.

(3) Show *ALL* your work on the exam pages.

(4) Do not discuss or provide information to anyone concerning the questions on this exam or your solutions until I post the solutions to the exam.

(5) You may use the following:

- Calculator - Your device cannot facilitate a connection to the internet or to send text messages
- Summary Sheets - (**8-pages,** 8.5"**x**11"**, write on both sides of the eight sheets**)
- The attached 22 pages of tables

(6) Do not use any other written material except for your summary sheets and the attachments to the exam.

(7) Do not use a computer, cell phone, or any other electronic device (other than a calculator).

I attest that I spent no more than 2 hours to complete the exam. I used only the materials described above. I did not receive assistance from anyone during the taking of this exam.

Student's Signature________________________________

I. (80 points - 4 points each) PLACE ONE of the following letters **(A, B, C, D, or E)** corresponding to the **BEST** answer on the **ANSWER SHEET - Page 10** of the Exam. I will also give partial credit for any work you display on the test for each of the problems.

(1.) An environmental group conducted a study of the impact of hydraulic fracking on pollution in ground water. The depth of the water table in areas surrounding the oil wells using fracking is an important factor in the level of pollution. This issue was addressed by classifying oil wells into 5 groups relative to the depth of the water table near the well. From each of the 5 groups of wells, a random sample of 20 oil wells was taken. At each of the 100 oil wells, water samples were taken at 25 randomly selected locations within a 1 mile radius of each well to determine the level of pollution in the ground water. This study is an example of

A. a simple random sample.

B. a simple random cluster sample.

C. a stratified simple random sample.

D. a stratified multistage cluster random sample.

E. a multistage cluster random sample.

(2.) A research project is studying the distribution of the number of mutations in 230 genes which may be associated with autistism. The researchers randomly selected 2000 autistic children and determined for each child how many of the 230 genes were mutated. Genetic theory states that mutations are independent from gene to gene and occur with approximately the same frequency. A probability model for M, the number of mutations in the 230 specified genes of an autistic child, is

A. Binomial

B. Geometric

C. Hypergeometric

D. Negative Binomial

E. Poisson

(3.) A metallurgist designed a study to estimate the distribution of cracks in the cooling pipes at nuclear power plants. She randomly selected 200 sections of pipes within the cooling systems at various nuclear power plants. An x-ray is taken of each pipe and the number of cracks, C, of length greater than 20mm is recorded. A possible probability model for C, the number of cracks of length greater than 20mm in a randomly selected pipe, is

A. Binomial

B. Negative Binomial

C. Weibull

D. Hypergeometric

E. Poisson

(4.) Which of the following statements about population parameters when both μ and σ exist is **FALSE**?

A. For highly right skewed distributions, the median is a better representative of the distribution "center" than the population mean μ.

B. For highly right skewed distributions, MAD is a better measure of variation than the population variance σ^2.

C. The population median $\tilde{\mu}$ can always be related to the population mean μ by $|\tilde{\mu} - \mu| \leq \sigma$.

D. A 5%-trimmed mean averages the middle 90% of the population values.

E. All the above statements are true.

(5.) A study is to be conducted to estimate the mean conductivity (in ohms) of a new alloy. Determine the sample size needed to ensure that the sample mean will estimate the true average conductivity to within 5 ohms with a reliability of 99%. From previous studies, conductivity had approximately a normal distribution with 99.7% of its values between 2 and 92 ohms.

A. 15

B. 35

C. 60

D. 135

E. cannot be determined with the given information

(6.) A researcher is studying the accuracy of gasoline station pumping devices. She wants to construct a 99% confidence interval for the average error in the pumping device using a random sample of 35 stations and measuring their errors: $Y_1, \ldots, Y_{35}$. A plot of the data reveals that the data is highly right skewed. The Shapiro-Wilk statistic for the transformation $X_i = log(Y_i)$ yields p-value=0.538. Which of the following methods would you recommend for constructing the confidence interval for the average error, μ_Y?

A. Use a distribution-free confidence interval because $n = 35$ is not very large.

B. Use the inverse of the end points of the confidence interval for μ_X, that is, use $\left(e^{L_X}, e^{U_X}\right)$.

C. Use a studentized Bootstrap confidence interval.

D. Use the standard t-Based confidence interval using the $Y-$data: $\bar{Y} \pm t_{.005,34} S/\sqrt{35}$, because $n > 30$.

E. None of the above procedures would be acceptable.

(7.) A study was designed to compare the mean yield of a genetically modified variety of broccoli to the most widely cultivated variety of broccoli. In 100 one acre plots of land, the genetically modified variety of broccoli was raised and the annual broccoli yields were recorded: $B_1,\ Y_2,\ \ldots,\ B_{100}$. From USDA records, the broccoli yields per acre of 103,259 farms were recorded: $Y_1,\ Y_2,\ \ldots,\ Y_{103259}$. The sample means $\bar{B}$ and $\bar{Y}$ are computed from the two samples. Which of the following statements is **TRUE** about the bias of $\bar{B}$ and $\bar{Y}$ as estimators of μ_B and μ_Y, respectively?

A. $\bar{B}$ and $\bar{Y}$ have the same positive bias

B. $\bar{B}$ has a larger bias than $\bar{Y}$

C. $\bar{B}$ has a smaller bias than $\bar{Y}$

D. the estimator having largest bias depends on the shape of the population pdf's

E. none of the above statements are true

(8.) A quality control officer for a manufacturer of steel beams randomly selects 29 beams from a warehouse of 1000 beams and measures the tensile strength of the 29 beams. A 95% confidence interval for the mean tensile strength of the beams in the warehouse is constructed to be (12.4, 23.6). Which of the following statements would be the most appropriate interpretation of this interval?

A. The mean strength of the beams in the warehouse has a 95% probability of falling between 12.4 and 23.6.

B. There is a 95% probability that the sample mean of any sample of 29 beams will be between 12.4 and 23.6.

C. Approximately 95% of all sample means for 29 beams are between 12.4 and 23.6.

D. If in the future a random sample of 29 beams is selected from the warehouse then there is a 95% chance that the mean of the 29 beams will be between 12.4 and 23.6.

E. None of the above are correct interpretations.

(9.) The significance probability, p-value, of the computed value of a test statistic is

A. the probability of observing a value of the test statistic less extreme to H_o

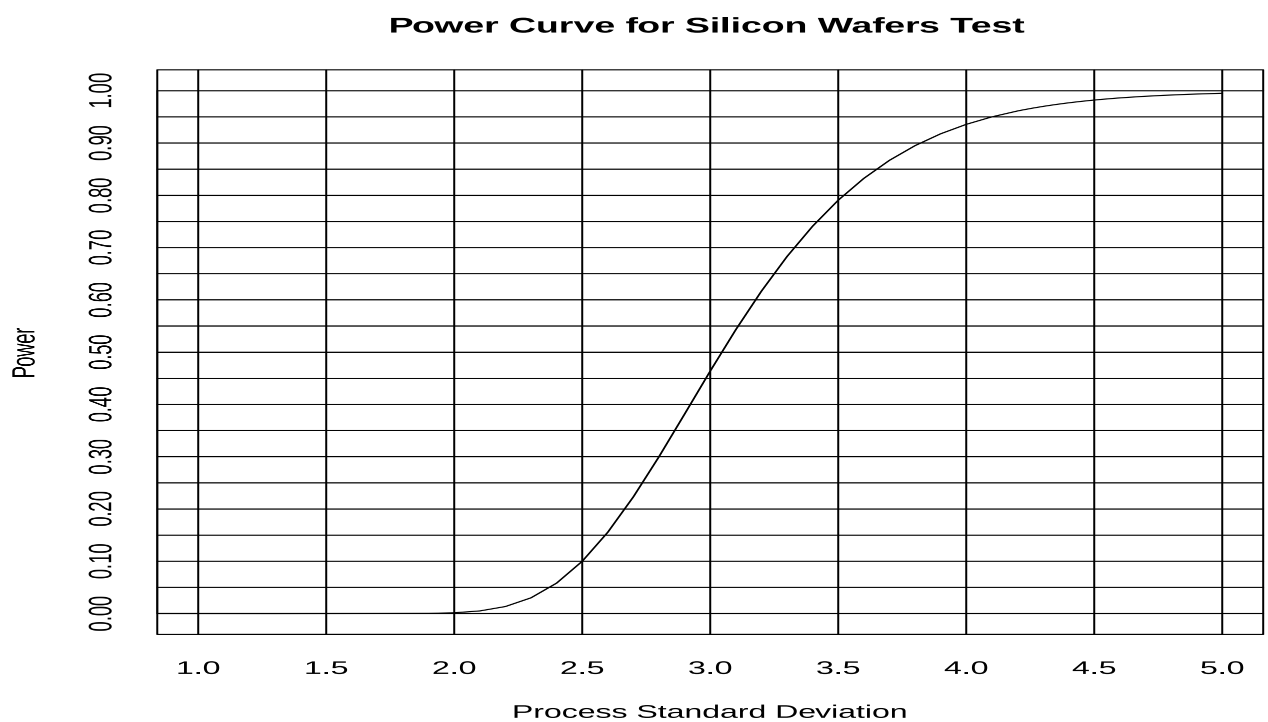
B. the weight of evidence in favor of H_1

C. the largest value of α for which the observed data will reject H_o

D. the smallest value of α for which the observed data will reject H_o

E. none of the above are appropriate descriptions of the p-value

(10.) A company that manufacturers silicon wafers for computer chips is concerned with both the mean thickness of the chips and the fluctuation in the thickness of the chips. In order to monitor the thickness, a random sample of 10 chips is selected every hour and the thickness is measured on each of the chips. The process is considered to be in control provided the process mean, μ, is 200 mm and the process standard deviation, σ is less than or equal to 2.5 mm. The company's process engineer develops a test to evaluate whether the process standard deviation is greater than 2.5 mm. She plots the power curve of the test in order to evaluate its performance. The curve is given here:



Using the power curve depicted above, what is the maximum probability of a Type I error and what is the probability of a Type II error if the population $\sigma = 3.5$?

A. .05 and .95

B. .10 and .21

C. .10 and .79

D. .05 and .79

E. cannot be determined from the power curve

(11.) An experiment is run to study the effects of DDT, a banned pesticide, on the bladder thickness of field mice. A Department of Natural Resources (DNR) researcher measured the bladder thickness of 8 mice exposed to DDT. After a 6 months exposure to DDT, the bladder thickness of the 8 mice was remeasured. From previous studies, the distribution of bladder thicknesses has been well fit by a double exponential distribution.

Mouse	1	2	3	4	5	6	7	8
PreExposure:	0.6	6.2	0.8	6.0	6.8	0.8	0.2	0.4
PostExposure:	3.3	3.4	3.6	3.5	3.4	3.5	3.4	3.6

Which test statistics would be most appropriate for determining if there is significant evidence that the average bladder thickness is smaller after exposure to DDT?

A. pooled t-Test

B. paired t-Test

C. Welch-Satterthwaite Separate Variance t-test

D. Wilcoxon Signed Rank test

E. Wilcoxon Rank Sum test

(12.) A metallurgist for a steel company is investigation the strength of a new alloy at various times after the alloy has been heat treated. He independently heat treats 25 specimens of the alloy and records the tensile strength of the alloy every 24 hours for the 30 days following the heat treatment yielding 30 strength values for every specimen: $X_{1i}, \ldots, X_{30i}$ for each of the $i = 1, \ldots, 25$ specimens. He uses the mean responses from the 25 specimens, $\bar{X}_1, \ldots, \bar{X}_{25}$, to test the hypotheses $H_o : \mu \leq 2800$ versus $H_1 : \mu > 2800$, where μ is the average tensile strength of the alloy prior to heat treating. An evaluation of the data reveals the following:

- The Shapiro-Wilk test has p-value=0.312 for each of the 25 specimens
- For $i = 1, \ldots, 25$: $X_{1i}, \ldots, X_{30i}$ have a lag one autocorrelation values $\hat{\rho}_i$ satisfying $\min[\hat{\rho}_1, \cdots, \hat{\rho}_{25}] > 0.89$.

A test was constructed using the following decision rule: $\frac{\bar{\bar{X}}-2800}{S/\sqrt{25}} \geq t_{.05,24}$, where $\bar{\bar{X}}$ and S are the mean and standard deviation of $\bar{X}_1, \ldots, \bar{X}_{25}$. The true level of significance of this test would be

A. very close to 0.05.

B. much less than 0.05.

C. much greater than 0.05.

D. may be greater or less than 0.05 depending on the value of σ.

E. can not be determined with the given information

(13.) A random sample of 27 units was taken from a population yielding values $Y_1, \ldots, Y_{27}$. From this data, the following confidence intervals for the population mean μ were computed: $\bar{Y} \pm t_{\alpha/2,26} S/\sqrt{27}$

90% $C.I.$ $(17.5, 24.5)$; 95% $C.I.$ $(16.8, 25.2)$; 99% $C.I.$ $(15.3, 26.7)$

In testing the hypotheses: $H_o : \mu \geq 25$ versus $H_1 : \mu < 25$, the $p-value$ computed from the data is

A. p-value $< .005$

B. $.005 <$ p-value $< .01$

C. $.01 <$ p-value $< .025$

D. $.025 <$ p-value $< .05$

E. $.05 <$ p-value

(14.) An experiment is conducted to test the research hypothesis $H_1 : \theta_1 < \theta_2$ where θ_1 and θ_2 are the location parameters from two populations. The data consists of random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ with the $X_i's$ independent of the $Y_i's$. Box plots of the data reveals the following:

- The box plots of both the X_i's and Y_i's have more than 20% of the data values identified as outliers
- The two box plots are nearly identical in shape but the $X-$values are generally larger than the $Y-$values.

The preferred test statistic is

A. Wilcoxon Rank Sum test

B. Wilcoxon signed rank test

C. Pooled t-test

D. Separate variance t-test

E. Sign test

(15.) An engineer for an automotive manufacturer is studying the occurrence of a defective in the braking system for a newly designed braking system. She randomly selects 100 cars installed with the braking system and the cars are placed into service at a car rental agency. The study terminates after the 75th automobile incurs a brake failure. For each of the 100 automobiles she recorded the mileage at which a failure occurred in the braking system or the mileage driven until the study is terminated. We would describe the data from this type of study as having

A. Type I censored data

B. Type II censored data

C. Random censored data

D. Right censored data

E. Uncensored data

(16.) A safety inspector of the National Highway Bureau is investigating the viability of bridges throughout the United States. In particular, she wants to determine if the design of the support beams in the bridges are related to major cracking in the support beams. Nearly all bridges have one of 6 designs used in their construction of bridge supports. For each of the 6 designs, she randomly selects 150 bridges using that design. The 900 bridges are then inspected and the number of major cracks in their support beams are recorded. What test procedure would best answer the safety inspector's question about the relationship between the six bridge support designs and the average number of major cracks per bridge?

A. Bonferroni adjusted pooled t-Test

B. Bonferroni adjusted Wilcoxon Rank Sum Test

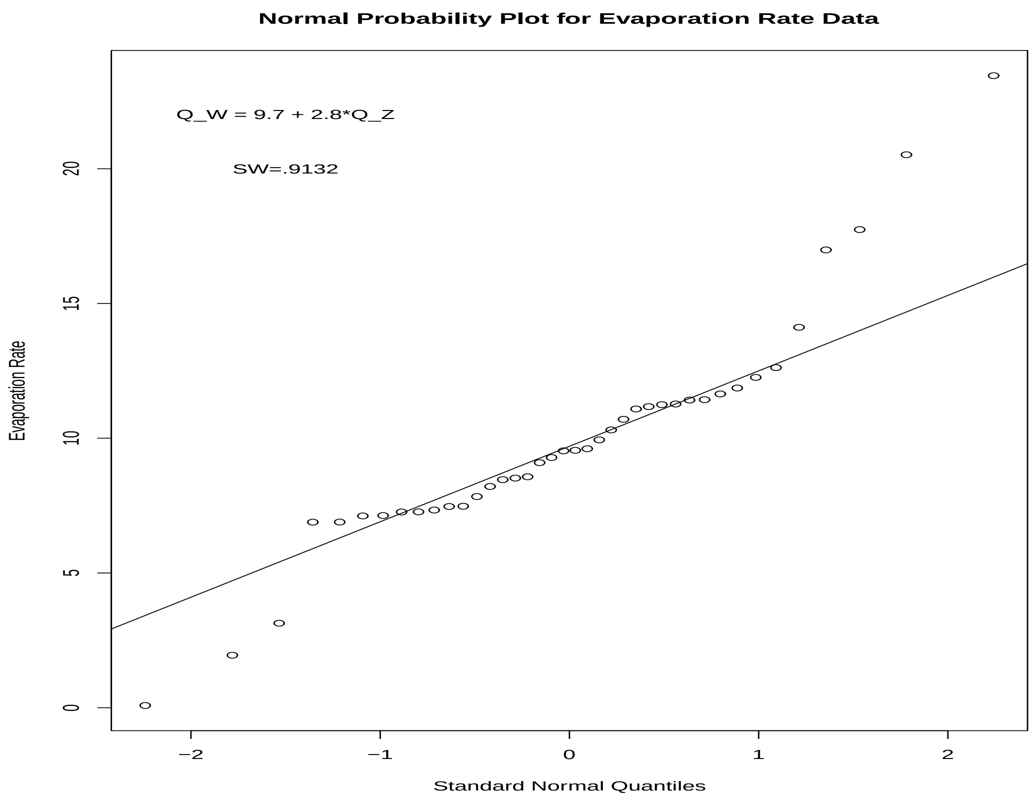
C. Chi-square test of homogeneity of proportions

D. Hartley's F-max test

E. Breslow-Day test of homogeneity of odds ratios

The following discussion will supply the information for Questions 17-18. A meteorologist is studying the effect of cloud cover on the reduction in surface water evaporation. She knows that the median evaporation rate in an environment having coastal humidity levels in July is 10 units per hour. She wants to determine if the median evaporation rate is less than 10 units per hour in a region having a very cloudy environment. Forty regions having cloudy July days were used in the experiment and their evaporation rates at midday were recorded.

(17.) The meteorologist conducted the experiment and the 40 evaporation rates produced the following normal probability plot with plotted line $\widehat{Q}(u) = 9.7 + 2.8 * Q_Z(u)$ and a value of .9132 from the Shapiro-Wilk test.



Which of the following distributions provides the best fit to the data?

A. normal

B. Weibull

C. t with df=23

D. Cauchy

E. Uniform on (0, 25)

(18.) For the water evaporation data, the power curve for the Sign Test having the same nominal size as the t-Test would be

A. below the power curve for the t-test

B. above the power curve for the t-test

C. very close to the power curve for the t-test because $n = 40$ is a large sample size

D. below the t-test power curve for $\tilde{\mu} < 10$ units

E. above the t-test power curve for $\tilde{\mu} < 10$ units

Using the following information to answer Questions 19-20. Preference for national income tax reform is often associated with political party. To estimate the preference for reform, a random sample of 4514 registered voters was obtained. The selected voters were asked their opinion on tax reform and what party they were affiliated with. The voters were also classified according to their annual family income.

	Party		
Tax Change	Rep.	Dem.	Total
Yes	772	1572	2344
No	983	1187	2170
Total	1755	2759	4514

	Income (thousands dollars)					
	20-59		60-99		Over 100	
	Party		Party		Party	
Tax Change	Rep.	Dem.	Rep.	Dem.	Rep.	Dem.
Yes	246	753	323	586	178	258
No	238	498	413	483	307	231
OR	.6836		.6446		.5191	

(19.) Which of the following procedures would provide the most valid assessment of a relationship between Tax Change Preference and Political Party if the Family Income of the voters is ignored?

A. Chisquared test of homogeneity of proportion

B. Chisquared test of independence

C. Breslow-Day test of homogeneity of odds ratios

D. Cochran-Mentel-Haenszel test of association

E. Brown-Forsythe-Levene test of homogenity

(20.) Which of the following procedures would provide the most valid assessment of a relationship between Tax Change Preference and Political Party if the Family Income of the voters is taken into account?

A. Chisquared test of homogeneity of proportion

B. Chisquared test of independence

C. Breslow-Day test of homogeneity of odds ratios

D. Cochran-Mentel-Haenszel test of association

E. Brown-Forsythe-Levene test of homogenity

II. (20 points) Show all the steps in your solutions to the following three problems.

(A.) A tire company has developed a new tread design. To determine if tires with the new design have a mean life greater than tires with the standard design, a random sample of n prototype tires and n standard tires will be tested. From previous studies, it is known that the distribution of tire life has approximately a normal distribution with $\sigma = 3,000$ miles. Assume that the distribution of the prototype tires is also approximately normal with $\sigma = 3,000$ but with possibly a larger mean. Determine the appropriate value of n such that an $\alpha = .05$ test will have power of at least 90% to detect an increase of 1500 miles in the mean tire life of the prototype tires over the mean life of standard designed tires.

(B.) When a milk producer receives complaints from its customers that the producer's milk tends to spoil prematurely, the producer needs to confirm whether the problem arises prior to shipping the milk. Milk spoilage is measured by the number of bacteria that are found in milk after it has been stored for 2 days. The average bacteria count should be 0.9 SPC's or less in order to meet USDA standards. The producer randomly selects 10 milk shipments for inspection and records the bacteria counts for the 10 shipments: $C_1, \ldots, C_{10}$. The values are given here:

Shipment	1	2	3	4	5	6	7	8	9	10	Total
Bacteria Count	1	3	1	2	1	3	2	1	3	2	19

You may assume that the 10 bacteria counts $C_1, \ldots, C_{10}$ are *iid* realizations from a Poisson random variable with an average bacteria count of λ and that $T = \sum_{i=1}^{10} C_i$ has a Poisson distribution with parameter 10λ

(1.) Compute the p-value for testing if the average bacteria count is greater than 0.9 SPC. Is there is significant evidence at the $\alpha = .011$ level that the average bacteria count is greater than 0.9 SPC? (Hint: Use Table A.2 in provided tables.)

(2.) Compute the probability that the test derived in part (1.) would fail to detect that the true average bacteria count is 2.0 SPC's.

ANSWER SHEET FOR MULTIPLE CHOICE QUESTIONS

Name ____

Place your answer to each of the MULTIPLE CHOICE questions on this page. Place only one of the UPPER case letters: A, B, C, D, E for each question.

(1.)________ (11.)________

(2.)________ (12.)________

(3.)________ (13.)________

(4.)________ (14.)________

(5.)________ (15.)________

(6.)________ (16.)________

(7.)________ (17.)________

(8.)________ (18.)________

(9.)________ (19.)________

(10.)________ (20.)________

STATISTICS 641 - FINAL EXAMINATION - SOLUTIONS

I. (80 points - 4 points each) SELECT ONE of the following letters **(A, B, C, D, or E)**

(1.) Correct Answer - D - Strata = 5 depth levels, clusters = wells, wells are clusters of water samples, random sample of 20 wells within each stratum/depth, random sample of 25 water samples within each well

(2.) Correct Answer - A - M is number of successes in 230 iid Bernoulli trials with M = number of mutations

(3.) Correct Answer - E - This cannot be binomial because the number of possible cracks per pipe is random, not a value from 0 to a fixed number.

(4.) Correct Answer - E - All the statements were true.

(5.) Correct Answer - C - $6\sigma \approx (92 - 2) \;\Rightarrow\; \sigma \approx 15 \;\Rightarrow\; n = \frac{\sigma^2 (Z_{.01/2})^2}{D^2} = \frac{(15)^2(2.58)^2}{(5)^2} = 59.9$

(6.) Correct Answer - C - log-normal distribution is right skewed and transformations are not appropriate for C.I.s for μ and σ.

(7.) Correct Answer - E - $\bar{B}$ and $\bar{Y}$ are both unbiased estimators of the mean

(8.) Correct Answer - E - The correct interpretation is that in a very large number of 95% C.I.'s for the population mean approximately 95% of the C.I.'s will contain the population mean and approximately 5% of the C.I.'s will fail to contain the population mean.

(9.) Correct Answer - D - Reject H_o if p-value $\leq \alpha$. If p-value is larger than α, H_o is not rejected. Thus, if a value of α smaller than the p-values is selected, then the null hypothesis would not be rejected.

(10.) Correct Answer - B - Max P(Type I error occurs at $\sigma = 2.5$) = Power($\sigma = 2.5$)=.10 and P(Type II error at $\sigma = 3.5$) =1- Power($\sigma = 3.5$) = 1-.79 = .21

(11.) Correct Answer - D - The data is paired and the distributions are heavy-tailed

(12.) Correct Answer - A - Although the 30 observations on each of the 25 specimens are highly correlated the 25 specimens have **independent** sample means, $\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_{25}$. Therefore, the test will maintain its specified level of significance because it is based on the means not the individual observations.

(13.) Correct Answer - D - The p-value associated with testing $H_o : \mu \geq 25$ versus $H_1 : \mu < 25$ is obtained as follows:

$\mu = 25$ is not in the 90% C.I. which implies reject H_o hence $p - value < \frac{1-.9}{2} = .05$

$\mu = 25$ is in the 95% C.I. which implies fail to reject H_o hence $p - value > \frac{1-.95}{2} = .025$

Thus, we have that $.025 < p - value < .05$

(14.) Correct Answer - E - Comparing two distributions using two independent random samples from two very heavy tailed distributions

(15.) Correct Answer - D - The number of miles until failure for the censored cars is greater than the recorded value. It is not Type II because the number of miles traveled by the censored cars could be more than or less than the number of miles traveled for the 75th automobile incurring a brake failure.

(16.) Correct Answer - B - The distributions of Number of Cracks for the six bridge designs would not have a normal distribution

(17.) Correct Answer - D - The distribution is symmetric with both right and left tails much heavier than a normal distribution.

(18.) Correct Answer - E - the distribution is very heavy tailed, the Sign test will have greater power than the t-test for values of the parameter in the Alternative Hypothesis. For values of the parameter in the Null Hypotheses, the power curve for the Sign test will be equal to or below the power curve for the t-test.

(19.) Correct Answer - B - Testing for Independence between Tax Preference and Political Party ignoring Family Income

(20.) Correct Answer - D - In order to take into account the effect of Family Income, it is necessary to use the adjustments provided by the CMH procedure.

II. (20 points) Show all the steps in your solutions to the following three problems.

(A.) This is a two-sample problem hence $n = \frac{2(Z_{.05}+Z_{.1})^2\hat{\sigma}^2}{\delta^2} = \frac{2(1.645+1.28)^2(3000)^2}{(1500)^2} = 68.4 \Rightarrow$ Need at least 69 tires of each design.

(B.) This problem is an extension of the exact binomial tests as given on pages 56-57 in Handout 12. This problem just replaces binomial with Poisson.

(1.) Test the hypotheses: $H_o : \lambda \leq 9$ vs $H_1 : \lambda > 9$ using $T = \sum_{i=1}^{10} C_i$ which has a Poisson distribution with parameter $\lambda_T = 10\lambda$. Reject H_o for a large value of T:

From the data, $T = 19$, thus p-value $= P[T \geq 19$ with $\lambda = .9] = 1 - P[T \leq 18$ with $\lambda = .9] = 1 - .998 = .002$ using Table A.2 with $\lambda_T = 10(.9) = 9$

$p - value = .002 < .011 = \alpha$ which implies reject H_o and hence reject H_o and conclude there is significance evidence that the bacteria count is greater than 0.9 SPC.

(2.) Reject H_o if $T \geq T_{.011,\lambda=9} = 17$ because $P[T \geq 17] = 1 - P[T < 17] = 1 - P[T \leq 16] = 1 - .989 = .011$ using Table A.2 with $\lambda_T = 10(.9) = 9$

$$\begin{aligned} P\,[\text{Type II error at } \lambda = 2] &= P\,[\text{Fail to reject} H_o \text{ at } \lambda = 2] \\ &= P\,[T < 17 \text{ at } \lambda = 2] \\ &= P\,[T \leq 16 \text{ at } \lambda = 2] \\ &= .221 \text{ using Table A.2 with } \lambda_T = 10(2) = 20 \end{aligned}$$

- Several students attempt to use the CLThm to use a Z statisics as the test statistic but n=10 is too small to apply an asymptotic approach.
- If n was larger the appropriate Z test would have been as follows:

Reject H_o if $Z = \dfrac{\bar{C} - .9}{\sqrt{.9}/\sqrt{10}} \geq Z_{.011} = 2.29 \;\Rightarrow\; p{-}value = P(Z \geq \dfrac{1.9 - .9}{\sqrt{.9}/\sqrt{10}} = P(Z \geq 3.33) = .0004 < .011$

- Thus, reject H_o and conclude there is significance evidence that the bacteria count is greater than 0.9 SPC.
- P(Type II error if $\lambda = 2$) $= P(\frac{\bar{C}-.9}{\sqrt{.9}/\sqrt{10}} \geq 2.29$ when $\lambda = 2) = P(Z \geq \frac{.9-2}{\sqrt{2}/\sqrt{10}} + 2.29\sqrt{.9/2}) = .1788$

<u>**EXAM I SCORES:**</u> n = 70

Min = 24, Q(.25) = 57, Q(.5) = 64 , Mean = 64.5, Q(.75) = 73, Max = 92