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