MAT 2377C
Final exam

December 10, 2010 **Professor: Rafal Kulik**
Time: 180 minutes

Student Number: __________________

Family Name: __________________

First Name: __________________

This is an open book examination.
Only non-programmable and non-graphic calculators are permitted.
Record your answer to each question in the table below.
Number of pages: **7** (including this one).
Number of questions: **24**.
NOTE: At the end of the examination, hand in only this page. You may keep the questionnaire.

Question	Answer	Question	Answer
1		13	
2		14	
3		15	
4		16	
5		17	
6		18	
7		19	
8		20	
9		21	
10		22	
11		23	
12		24	

Q1. Suppose that for a very large shipment of integrated-circuit chips, the probability of failure for any one chip is 0.09. Find the probability that at most 2 chips fail in a random sample of size 20. (The numbers are rounded down to the second decimal place).

(a) 0.13 (b) 0.90 (c) 0.73 (d) 0.20 (e) none of the preceding

Solution to Q1:

$X \sim B(20, 0.09)$. So

$$P(X \leq 2) = \binom{20}{0}(.09)^0(.91)^{20} + \binom{20}{1}(.09)^1(.91)^{19} + \binom{20}{2}(.09)^2(.91)^{18} = 0.73$$

Q2. A scientist inoculates several mice, one at a time, with a disease germ until he finds one that has contracted the disease. If the probability of contracting the disease is 1/6, what is the probability that 8 mice are required? (The numbers are rounded down to the fourth decimal place).

(a) 0.0465 (b) 0.9350 (c) 0.0740 (d) 0.2600 (e) none of the preceding

Solution to Q2:

$X = \#$ number of mice required to observe the first one with the disease. $X \sim$ is geometric with $p = 1/6$. So $P(X = 8) = (1-p)^7\, p = 0.0465$

Q3. Assume that random variables X and Y are independent and have distribution $X \sim \text{Poisson}(2)$ and $Y \sim B(5, 0.2)$, respectively. Compute

$$P(\{X = 0\} \text{ or } \{Y = 0\}) = P(\{X = 0\} \cup \{Y = 0\}) \; .$$

(The numbers are rounded up to the fourth decimal place)

(a) 0.4325 (b) 0.5621 (c) 0.4630

(d) 0.4187 (e) insufficient information provided

Solution to Q3:

$$\begin{aligned} P(\{X = 0\} \text{ or } \{Y = 0\}) &= P(\{X = 0\}) + P(\{Y = 0\}) - P(\{X = 0\} \text{ and } \{Y = 0\}) \\ &= P(\{X = 0\}) + P(\{Y = 0\}) - P(\{X = 0\})\, P(\{Y = 0\}) \\ &= \exp(-2) + 0.2^0 * 0.8^5 - \exp(-2) * 0.2^0 * 0.8^5 = 0.4187 \end{aligned}$$

Q4. A medical research team wished to evaluate a proposed screening test for Alzheimer's disease. The test was given to a random sample of 450 patients with Alzheimer's disease, in 436 cases the test result was positive. Also, the test was given to a random sample of 500 patients without the disease, only in 5 cases the result was positive. It is known that in the Canada 7.7% of the population aged 65 and over have Alzheimer's disease.

Find the probability that a person has the disease given that the test was positive. (The numbers are rounded down to the second decimal place).

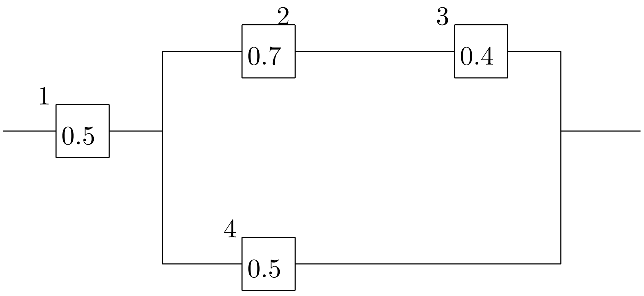
(a) 0.97 (b) 0.88 (c) 0.99 (d) 0.12 (e) none of the preceding

Solution to Q4:

A - test positive, D - a person has disease. Given: $P(A|D) = \frac{436}{450}$, $P(A|D^c) = \frac{5}{500}$, $P(D) = 0.077$. To find: $P(D^c|A)$ (Bayes' formula):

$$P(D|A) = \frac{P(A|D)P(D)}{P(A|D)P(D) + P(A|D^c)P(D^c)} = 0.88.$$

Q5. Consider the following system with four components. We say that it is functional if there exists a path of functional components from left to right. The probability of each component functions is shown. Assume that the components function or fail independently. What is the probability that the system operates? (The numbers are rounded down to the second decimal place).



(a) 0.32 (b) 0.16 (c) 0.03 (d) 0.68 (e) none of the preceding

Solution to Q5:

Call 'Box B' - components 2,3,4, 'Box C' - components 2,3.

$$\begin{aligned} P(\text{Box C operates}) &= P(\text{component 2 operates and component 3 operates}) \\ &= P(\text{component 2 operates})P(\text{component 3 operates}) = 0.4 \times 0.7 = 0.28. \end{aligned}$$

$$\begin{aligned} P(\text{Box B operates}) &= P(\text{Box C operates or component 4 operates}) \\ &= P(\text{Box C operates}) + P(\text{component 4 operates}) - \\ &\quad P(\text{Box C operates})P(\text{component 4 operates}) \\ &= 0.28 + 0.5 - 0.28 * 0.5 = 0.64. \end{aligned}$$

$$\begin{aligned} P(\text{system operates}) &= P(\text{component 1 and Box B operate}) \\ &= P(\text{component 1 operates})P(\text{Box B operates}) \\ &= 0.5 * 0.64 = 0.32. \end{aligned}$$

Q6. A material is studied for a possible contamination. Suppose that occurrences of contaminated particles can be described by a Poisson process with the intensity 0.02 particles per kilogram. Find the probability that there will be no contaminated particles in 10 kilograms of the material.

(a) 0.989 (b) 0.819 (c) 0.135 (d) 0.020 (e) none of the preceding

Q7. In a NiCd battery, a fully charged cell is composed of Nickelic Hydroxyde. Nickel is an element that has a multiple oxidation states. Let X be the nickel charge, which has the following probability mass function:

x	$f_X(x)$
0	.18
2	.34
4	.33
10	k

where k is an unique number. Determine the mean and the standard deviation of the nickel charge. (The numbers are rounded down to the third decimal place).

(a) 3.064; 3.5 (b) 3.5; 9.390 (c) 3.5; 3.064 (d) 4; 7.250 (e) none of the preceding

Q8. Let X be a continuous random variable with mean $\mu_X = .75$ and probability density function

$$f_X(x) = 3\,x^2,\ 0 < x < 1.$$

Give the variance of the random variable X. (The numbers are rounded down to the fourth decimal place).

(a) 0.0375 (b) 0.1936 (c) 0.7746 (d) 0.3873 (e) none of the preceding

Solution to Q8:

$$\mathrm{E}[X^2] = \int_0^1 x^2\,(3\,x^2)\,dx = 0.6 \ \Rightarrow\ \sigma_X^2 = E[X^2] - \mu_X^2 = 0.0375$$

Q9. Assume that X_i, $i = 1, \ldots, 81$, are independent random variables with a common density function $f_X(x) = \exp(-x)$, $x > 0$. Find the approximate probability $P\left(\sum_{i=1}^{81} X_i > 90\right)$. Hint: Central Limit Theorem. (The numbers are rounded down to three decimal place).

(a) 0.644 (b) 0.168 (c) 0.346

(d) 0.158 (e) none of the preceding

Solution to Q9:

$\mathrm{E}(X) = 1$, $\mathrm{Var}(X) = 1$. So

$$P(\sum_{i=1}^{n} X_i > 90) \quad = \quad P\left(\frac{\sum_{i=1}^{81} X_i - 81 \times 1}{\sqrt{81 \times 1}} > \frac{90 - 81}{\sqrt{81}}\right) = P(Z_0 > 1) = 1 - 0.67 = 0.1586$$

Q10. A random sample of size $n_1 = 16$ is selected from a normal population with a mean of 75 and variance of 288. A second random sample of size $n_2 = 9$ is taken independently from another normal population with mean 80 and variance of 162. Let $\bar{X}_1$ and $\bar{X}_2$ be the two sample means. Find the probability that $\bar{X}_1 + \bar{X}_2$ exceeds 156.5. (The numbers are rounded down to the fourth decimal place).

(a) 0.5987 (b) 0.4012 (c) 0.6231 (d) 0.4235 (e) none of the preceding

Solution to Q10:

$\overline{X}_1 + \overline{X}_2 \sim N(75 + 80, 288/16 + 162/9) = N(155, 36)$.

$$P(\overline{X}_1 + \overline{X}_2 > 156.5) = P(Z > \frac{156.5 - 155}{\sqrt{36}}) = 1 - \Phi(0.027) = 0.4012.$$

Q11. The city of Ottawa would like to know how many people are in favor of an increase in a property tax. 1000 residents have been asked and 300 answered "Yes, I favor a property tax increase". A 95% confidence interval for the proportion of people, who are in favor of property tax increase, is

(a) [0.202,0.254] (b) [0.197,0.259] (c) [0.194,0.262]

(d) [0.272,0.328] (e) none of the preceding

(The lower limit was rounded up to the third decimal place; The upper limit was rounded down to the third decimal place).

Solution to Q11:

$$\widehat{p} \pm z_{.025}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} = \frac{300}{1000} \pm 1.96\sqrt{\frac{300/1000\,(1 - 300/1000)}{1000}}$$

So $[0.272, 0.328]$.

Q12. A machine is producing metal pieces that are cylindrical in shape. A sample of pieces is taken and the diameters are

1.01	0.97	1.03	1.04	0.99
0.98	0.99	1.01	1.03	

The sample mean and standard deviation are $\overline{x} = 1.00556$ and $s = 0.02455$, respectively. Find a 95% confidence interval for the true mean diameter. Assume that the population is normally distributed.

(a) [0.989,1.022] (b) [0.987,1.024] (c) [0.978,1.033]

(d) [0.991,1.034] (e) none of the preceding.

(The lower limit was rounded up to the third decimal place; The upper limit was rounded down to the third decimal place).

Solution to Q12:

We have $t_{.025,9-1} = 2.306$. So

$$\overline{x} \pm 2.306\,\frac{s}{\sqrt{n}} = [0.987, 1.024].$$

Q13. Past experience has indicated that the breaking strength of yarn used in manufacturing drapery material is normally distributed and that $\sigma = 2$. A random sample of twenty five specimens is tested and the average breaking strength is found to be $\bar{x} = 98$ psi. Find a 99% confidence interval on the true mean breaking strength.

(a) [97.216,98.784] (b) [97.216,98.554] (c) [96.972,99.028]

(d) [97.446,98.554] (e) none of the preceding

(The lower limit was rounded up to the third decimal place; The upper limit was rounded down to the third decimal place).

Solution to Q13:

Conditions: normal population with known σ.
A 95% confidence interval is

$$\overline{x} \pm z_{.025} \frac{\sigma}{\sqrt{n}} = 98 \pm 1.96 \left(\frac{2}{\sqrt{25}} \right) = [97.216, 98.784].$$

Q14. We want to test the hypothesis that the average content of containers of a particular lubricant is <u>more</u> than 10 liters if the contents of a random sample of 10 containers are

10.2	9.7	10.1	10.5	10.1
11.1	9.9	10.4	10.4	9.5

Find the P-value of the one-sided test. Assume that the distribution of contents is normal. Hint: for this data we have

$$\sum_{i=1}^{10} x_i = 101.9 \text{ and } \sum_{i=1}^{10} x_i^2 = 1040.19.$$

(a) $0.05 < P < 0.10$ (b) $0.10 < P < 0.20$ (c) $0.25 < P < 0.40$

(d) $0.50 < P < 0.80$ (e) none of the preceding

Solution to Q14:

We test $H_0 : \mu = 10$ vs. $H_1 : \mu > 10$. We have $\overline{x} = 10.19$ and $s = 0.45$. The observed value of test statistics is

$$t_0 = \frac{\overline{x} - 10}{s/\sqrt{n}} = 1.332809$$

The P-value is $P(T > 1.332)$ and is between 0.1 and 0.2, see Table V with $\nu = n - 1 = 9$. Answer (B).

Q15. The melting point of each of 16 samples of certain brand of vegetable oil was determined, resulting in $\bar{x} = 94.32$. Assume that the distribution of melting point is normal with $\sigma = 1.2$. Test $H_0 : \mu = 95$ versus $H_1 : \mu \neq 95$ with $\alpha = 0.01$. Determine the P-value and state your conclusion.

(a) Reject H_0, P-value=0.012

(b) Can not reject H_0, P-value=0.023

(c) Reject H_0, P-value=0.023

(d) Can not reject H_0, P-value=0.012.

Solution to Q15:

The observed value of test statistics is

$$z_0 = \frac{\overline{x} - 95}{\sigma/\sqrt{n}} = \frac{94.32 - 95}{1.2/\sqrt{16}} = -2.27.$$

Since alternative is two-sided, P-value$= 2P(Z < -2.27) = 2\,(1 - \Phi(2.27)) = 0.023$. Since $P > \alpha$, we do not reject H_0.

Q16. An expert wishes to determine the average time (in seconds) that it takes to drill holes in a certain metal clamp. How large a sample is required to be 99% confident that the sample mean will be within 15 seconds of the true mean. Assume that it is known from previous studies that $\sigma = 30$ seconds.

(a) 26 (b) 27 (c) 28 (d) 29 (e) none of the preceding

Solution to Q16:

$$n \geq \left[\frac{z_{.005}\,\sigma}{E}\right]^2 = \left[\frac{(2.576)\,(30)}{15}\right]^2 = 26.5$$

So $n = 27$ observations.

Q17. Assume that we have a sample of size 10 from a population $N(4, 9)$. Denote by $\overline{X}$ and S^2, the sample mean and sample variance, respectively. Find c such that

$$P\left(\frac{\overline{X} - 4}{S/\sqrt{10}} \leq c\right) = .95.$$

(a) 1.833 (b) 1.86 (c) 1.645

(d) 2.262 (e) none of the preceding

Solution to Q17:

Equivalent statement: find c such that

$$P\left(\frac{\overline{X} - 4}{S/\sqrt{10}} \geq c\right) = 0.05.$$

We have that $\frac{\overline{X} - 4}{S/\sqrt{10}}$ has Student distribution with $n - 1 = 9$ degrees of freedom. From the table we read that $P(t_9 > 1.833) = 0.05$, thus $c = 1.833$.

Q18. The data below are from Darwin's study of cross- and self-fertilization. Pairs of seedlings of the same age, one produced by cross-fertilization and the other by self-fertilization, were grown together so that the members of each pair were treated under nearly identical conditions. The data are the final heights of each plant after a fixed period of time, in inches. Darwin consulted the famous 19th century statistician Francis Galton about the analysis of these data.

```
pair crossfertilized selffertilized
  1       23.5              17.4
  2       12.0              20.4
```

3	21.0	20.0
4	22.0	20.0
5	19.1	18.4
6	21.5	18.6
7	22.1	18.6
8	20.4	15.3
9	18.3	16.5
10	21.6	18.0
11	23.3	16.3
12	21.0	18.0
13	22.1	12.8
14	23.0	15.5
15	12.0	18.0

Assuming the heights are normally distributed, we conduct the test of no difference between fertilization method (two-sided alternative). Then

(a) P-value $\in (0.05, 0.1)$; do not reject H_0 for $\alpha = 0.01$;

(b) P-value $\in (0.05, 0.1)$; do reject H_0 for $\alpha = 0.01$;

(c) P-value $\in (0.025, 0.05)$; do not reject H_0 for $\alpha = 0.01$;

(d) P-value $\in (0.05, 0.1)$; do reject H_0 for $\alpha = 0.01$;

(e) none of the preceding

Solution to Q18:

This is paired test. Compute difference (crossfertilized-selffertilized). The obtained mean and standard deviation are $\bar{d} = 2.606667$, $s_D = 4.712819$. The test statistics is

$$T_0 = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{n-1}.$$

Its observed value is 2.142152. We compute P-value as $2P(t_{14} > 2.142152) \in 2 \times (0.025, 0.05) = (0.05, 0.1)$. Do not reject for $\alpha = 0.01$.

Q19. The thickness of a plastic film (in mils) on a substrate material is thought to be influenced by the temperature at which the coating is applied. A completely randomized experiment is carried out. Eleven substrates are coated at 125^oF, resulting in a sample mean coating thickness of $\bar{x}_1 = 103.5$ and a sample standard deviation of $s_1 = 10.2$. Another 11 substrates are coated at 150^oF, for which $\bar{x}_2 = 99.7$ and $s_2 = 11.7$ are observed. Assume normality. The value of the appropriate test statistics is (The numbers are rounded down to the third dp).

(a) 0.890; (b) 1.035; (c) 1.815; (d) 1.890;

(e) None of the preceding

Solution to Q19:

This is two-sample test. The observed value of the test statistics

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{103.5 - 99.7}{\sqrt{10.2^2/11 + 11.7^2/11}} = 0.8119597.$$

Q20. For a set of 12 pairs of observations on (x_i, y_i) from an experiment, the following summary for x and y is obtained: $\sum_{i=1}^{12} x_i = 25$, $\sum_{i=1}^{12} y_i = 432$, $\sum_{i=1}^{12} x_i^2 = 59$, $\sum_{i=1}^{12} y_i^2 = 15648$, $\sum_{i=1}^{12} x_i y_i = 880.5$. The estimated value of y at $x = 5$ from the least squares regression line is:

(a) 27.78 (b) 47.77 (c) 41.87 (d) 55.97 (e) none of the preceding

Q21. Assuming that the simple linear regression model $Y = \beta_0 + \beta_1 x + \varepsilon$ is appropriate from $n = 14$ observations we computed the estimated regression line $\widehat{y} = .6649 + .83075\,x$. Given that $S_{yy} = 4.1289$ and $S_{xy} = 4.4094$, compute the estimated standard error for the slope.

(a) 0.3176 (b) 0.0388 (c) 0.0855 (d) 0.0073 (e) none of the preceding

Q22. An engineer wants to study the variability of the production process. A sample of size $n = 10$ was taken every 2 hours. After $m = 25$ preliminary samples, one obtains $\bar{\bar{x}} = 30.2$ and $\bar{r}/d_2 = 2.5$. Determine the lower and the upper control limits for a R chart.

(a) LCL=1.716 and UCL=13.674
(b) LCL=4.511 and UCL=15.144
(c) LCL=27.830 and UCL=32.570
(d) LCL=28.696 and UCL=31.711
(e) none of the preceding

Solution to Q22:

From table XI, with $n = 10$, we obtain $d_2 = 3.078$, $D_3 = 0.223$ and $D_4 = 1.777$. So $\bar{r} = 2.5\,d_2 = 7.695$. The limits for R are

$$LCL = \bar{r}\,D_3 = (7.695)(0.223) = 1.716$$

et

$$UCL = \bar{r}\,D_4 = (7.695)(1.777) = 13.674.$$

Q23. A company manufactures computers. To control their quality, 50 computers are tested every day. We have the following data for the consecutive 15 days:

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
# of defective	3	2	5	1	4	1	6	3	5	0	6	2	4	1	7

The lower and upper control limits for the proportion of defective are, respectively:

(a) $LCL = -0.0391$, $UCL = 0.1725$
(b) $LCL = 0$, $UCL = 0.1725$
(c) $LCL = 0.0391$, $UCL = 0.1725$
(d) $LCL = -0.0391$, $UCL = 0.2599$

(e) None of the preceding.

Solution to Q23:

We have $\bar{p} = 0.67$, $n = 50$ so that

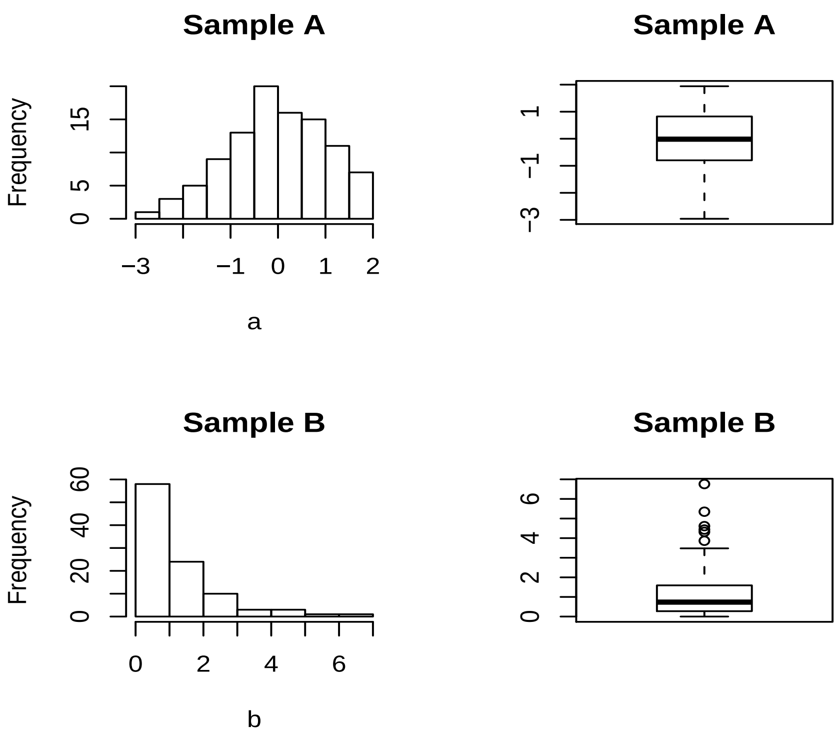
$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = -0.0391 \rightarrow LCL = 0,$$
$$UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0.1725$$

Q24. The following graphs show histogram and boxplot for two samples, A and B. Based on these graphs we may conclude that

(a) Only A comes from a normal population.

(b) Only B comes from a normal population.

(c) Both A and B come from normal populations.



This is the last question

Solutions to multiple choice questions:

Q1 $\longrightarrow$ c

Q2 $\longrightarrow$ a

Q3 $\longrightarrow$ d

Q4 $\longrightarrow$ b

Q5 $\longrightarrow$ a

Q6 $\longrightarrow$ b

Q7 $\longrightarrow$ c

Q8 $\longrightarrow$ a

Q9 $\longrightarrow$ d

Q10 $\longrightarrow$ b

Q11 $\longrightarrow$ d

Q12 $\longrightarrow$ b

Q13 $\longrightarrow$ a

Q14 $\longrightarrow$ b

Q15 $\longrightarrow$ b

Q16 $\longrightarrow$ b

Q17 $\longrightarrow$ a

Q18 $\longrightarrow$ a

Q19 $\longrightarrow$ e

Q20 $\longrightarrow$ a

Q21 $\longrightarrow$ c

Q22 $\longrightarrow$ a

Q23 $\longrightarrow$ b

Q24 $\longrightarrow$ a