SOLUTIONS

Family Name: _______________Given Name: ______________ Student Number: _________________

STATISTICS 2060/MATH 2060/ECON 2260 Final Exam
Friday Apr. 17, 2015

Please answer the questions in the space provided. Justify your answers.
A calculator, and three sheets of formulae/notes, letter size, are permitted.

1. Accidents on highway 103 occur according to a Poisson process at the rate of 1.5 per week.

(3) (a) What is the probability that there is at least one accident in a week?

$P(X \geq 1) = 1 - P(X=0) = 1 - e^{-1.5} = 1 - .2231 = .7769$

(3) (b) What is the probability that there are two accidents in the next two weeks? $Y \sim P(3)$

$$P(Y=2) = \frac{e^{-3} 3^2}{2!} = \frac{.0499 \times 9}{2} = .2240$$

(3) (c) What is the probability that there is exactly one accident in each of the next two weeks.

$P(X=1) = e^{-1.5} 1.5 = .3347 = p$

$\therefore \; p^2 = .3347^2 = .1120$

(3) (d) What is the probability that the first accident occurs within the first three days (3/7 week)?

$T \sim exp(\lambda = 1.5)$

$P(T \leq \frac{3}{7}) = 1 - e^{-1.5(3/7)}$

$= 1 - .5258$

$= .4742$

(3) 2. Suppose X is a random variable with mean 2 and variance 4, and Y is a random variable with mean 3 and variance 6, and that the covariance between X and Y is -1. What is the correlation between X and Y?

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{-1}{\sqrt{4 \times 6}} = \frac{-1}{2\sqrt{6}} = -.2041$$

3. The windshield wiper fluid reservoir for many automobiles has a 4 litre capacity. A randomly chosen car has wiper fluid, X, with the density function

$$f(x) = 1/16 + 3x/32$$

for $0 \leq x \leq 4$.

(3) (a) What is the cumulative distribution function for X?

$$F(x) = \int_0^x \left(\frac{1}{16} + \frac{3}{32}t\right)dt = \frac{1}{16}x + \frac{3}{64}x^2$$

(3) (b) What is the probability a car has more than 3 liters of fluid in its reservoir?

$$P(X > 3) = 1 - F(3) = 1 - \left(\frac{3}{16} + \frac{27}{64}\right)$$
$$= 1 - \frac{39}{64} = \frac{25}{64} = .3906$$

(4) (c) What is the median amount of wiper fluid in a car?

solve $\frac{1}{2} = \frac{1}{16}\eta + \frac{3}{64}\eta^2$

or $3\eta^2 + 4\eta - 32 = 0$

$$\eta = \frac{-4 \pm \sqrt{16 + 4(3)(32)}}{6} = \frac{-4 \pm 4\sqrt{1+24}}{6}$$
$$= \frac{-4 \pm 20}{6} \quad \text{use } (+) = \frac{16}{6} = \frac{8}{3} = 2.67$$

4. The letters of the word STATISTICS are put into a bag and ~~5~~ 4 letters are drawn without replacement.

(3) (a) What is the probability that the letters drawn spell the word CATS?

$$\frac{\binom{1}{1}\binom{1}{1}\binom{3}{1}\binom{3}{1}}{\binom{10}{4}} = \frac{9}{10\times9\times8\times7/4\times3\times2\times1}$$
$$= \frac{3}{70} = .0429$$

(3) (b) What is the probability that there are no Ts drawn?

$$\frac{\binom{3}{0}\binom{7}{4}}{\binom{10}{4}} = \frac{7\times6\times5/3\times2\times1}{10\times9\times8\times7/4\times3\times2\times1}$$
$$= \frac{6\times5\times4}{10\times9\times8} = \frac{6}{9\times4} = \frac{1}{6} = .1667$$

(3) 5. The random variables X_1, X_2 and X_3 are independent and identically distributed. Which has the larger variance, $Y = 3X_1$ or $W = X_1 + X_2 + X_3$? Justify your answer.

$V(Y) = 9V(X_1) \qquad V(W) = V(X_1) + V(X_2) + V(X_3) = 3V(X_1)$

$\therefore$ Y has the larger variance

6. Many car tires have a suggested air pressure of 220 kPa. Suppose the actual distribution of tire pressures is normal with mean 210 and standard deviation 15 kPa.

(4) (a) What is the probability that the difference between the pressure of two randomly chosen tires is greater than 30 kPa? $X-Y \sim N(0,\ 15^2+15^2 = 450)$

$$P(|X-Y|>30) = 2P(X-Y>30) = 2P\left(Z > \frac{30}{21.21}\right)$$
$$= 2P(Z<-1.41) = 2(.0793) = .1586$$

(4) (b) What is the probability that the mean tire pressure of 4 random tires is less than 210 kPa?

$$\bar{X} \sim N(210,\ 15^2/4 = 7.5^2)$$
$$P(\bar{X}<210) = P\left(Z < \frac{210-210}{7.5}\right) = P\left(Z< \frac{0}{7.5}\right)$$
$$= P(Z<0) = .50$$

7. A bivariate probability mass function is given below for the random variables X and Y related to whether car tires are underinflated and whether the tires are worn unevenly or not. $X=0$ if no tires are underinflated, $X=1$ if exactly one tire is underinflated, and $X=2$ if two or more tires are underinflated. $Y=0$ if no tires are worn unevenly and $Y=1$ if at least one tire is worn unevenly.

		X 0	1	2	Total
Y	0	.1	.1	.1	.3
	1	.1	.2	.4	.7
	Total	.2	.3	.5	

(3) (a) What is the conditional probability distribution for Y given that $X=1$? (List the possible values and their probabilities.)

y	0	1
$p(y\mid x=1)$	1/3	2/3

(3) (b) What is the covariance of X and Y? Use the fact that $E(X)=1.3$ and $E(Y)=.7$.

$$E(XY) = 1(1)(.2) + 1(2)(.4) = .2+.8 = 1.0$$
$$Cov(X,Y) = E(XY) - E(X)E(X) = 1-(1.3)(.7)$$
$$= .09$$

8. It is believed that the proportion of red-haired people in the city of Toronto is .0345. A random sample of 1000 Torontonians is selected. (You can assume sampling with replacement because the sample is small compared to the total population.) What is the approximate probability that fewer than 25 redheads are found?

(2) (a) State which approximation you are using, and explain why it is valid in this case.

Normal approx to binomial

$np = 34.5 \quad n(1-p) = 965.5$ both > 10

(4) (b) Give the mean and standard deviation of the number of redheads that will be found in the sample.

$\mu = np = 34.5 \quad \sigma^2 = npq = 34.5(1-.0345) = 33.31$

$\sigma = 5.77$

(4) (c) Approximate the probability that fewer than 25 redheads are found.

$$P(X < 25) \approx \Phi\left(\frac{25.5 - 34.5}{5.77}\right) = \Phi(-1.56) = .0594$$

or (better)

$$= P(X \leq 24) = \Phi\left(\frac{24.5 - 34.5}{5.77}\right) = \Phi(-1.73)$$

$$= .0418$$

9. A random sample of 10 digital photos taken outside in July have a mean size of 2.2 MB with a standard deviation of .3 MB. A random sample of 8 digital photos taken inside in January with the same camera have mean size 1.8 MB and standard deviation .4 MB.

(6) (a) Construct a 95% confidence interval for the mean size of July photos taken outside.

$\bar{y} \pm t_{.025, 9}(.3)/\sqrt{10}$

$2.2 \pm 2.262(.3)/3.16$

$\pm .215$

$\therefore (1.985, 2.415)$

(2) (b) What assumption(s) are required for this interval to be valid?

The sizes are a random sample from a normal distribution

(2) (c) Would a 90% confidence interval be narrower or wider?

Narrower

(d) Is the mean size of photos taken outside in July different from that of photos taken inside in January?

(2) i. State the hypotheses.

$H_0: \mu_1 = \mu_2 \quad (\mu_1 - \mu_2 = 0)$

$H_a \quad \mu_1 \neq \mu_2 \quad (\mu_1 - \mu_2 \neq 0)$

(5) ii. Calculate the test statistic.

$$s_p^2 = \frac{9(.3)^2 + 7(.4)^2}{16} = .1206 = (.3473)^2$$

$$t = \frac{2.2 - 1.8}{.3473\sqrt{\frac{1}{10} + \frac{1}{8}}} = \frac{.4}{.35(.4743)} = 2.428$$

(2) iii. Find the P value, using a bound if necessary. State the table used and show values from the table.

t - 16 df

$2.120 < 2.428 < 2.583$

so $.01 < P(T > t_{obs}) < .025$

and $.02 < P < .05$

(2) iv. Give a conclusion in the context of the problem.

We have strong evidence against the null hypothesis that the mean sizes are equal

(2) v. State what extra assumptions beyond those in part (b) are required for this test to be valid.

Both populations are normal.
The samples are independent
The variances are equal

(5) 10. A computer service center owner wishes to estimate the mean time it takes technicians to diagnose the problem with computers. How large a sample should she take to estimate the mean time if she believes the standard deviation of the diagnosis times is 10 minutes and she wishes a 95% confidence interval to the mean to have a margin of error (or bound on the error of estimation) of 5 minutes?

$z_{\alpha/2} = 1.96 \quad \sigma = 10 \quad \frac{w}{2} = 5$

$$n = \left(\frac{z_{\alpha/2}\,\sigma}{w/2}\right)^2$$

$$= \left(\frac{1.96 \times 10}{5}\right)^2 = 15.37$$

Round up $\Rightarrow n = 16$

11. In a recent study, subjects were shown two pictures of a person smiling. One of these pictures showed a real smile, while the other showed a fake smile. A random sample of 64 subjects found that 40 were able to correctly choose the real smile. Is this evidence that the subjects were able to distinquish the smiles better than by a random guess ($p = .5$)?

(2) (a) State the hypotheses.

$H_0 \; p = 1/2$

$H_a \; p > 1/2$

(4) (b) Calculate the test statistic.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{40/64 - .5}{\sqrt{.5(.5)/64}} = \frac{8/64}{\frac{1}{2}\frac{1}{8}} = 2.00$$

(2) (c) Calculate the significance probability or P value, as accurately as possible.

$P(Z > 2) = .0228$

(2) (d) Give a conclusion in terms of the strength of evidence against the null hypothesis.

There is very strong evidence against the null hypothesis that $p = 1/2$

(2) (e) Are the results statistically significant at the $\alpha = .05$ level of significance? Explain.

Yes, because $P < .05$

(2) (f) Describe briefly what would be a type II error in this case.

A type II error is failing to get significant evidence against H_0 when it is false. In this case this is getting a large P value for the H_0 $p = 1/2$ when in fact $p > 1/2$