Math 3081 **Final Exam Solutions** Summer 1 2018

1. (7 pts) 60% of the population received the flu vaccine last winter. The probability of catching the flu was 0.2 for a vaccinated person, and was 0.5 for an unvaccinated person. If a person did not catch the flu, what is the probability he had been vaccinated?

Let $V = vaccinated$ and $F = caught\ flu$

$$P(V|F') = \frac{P(V \cap F')}{P(F')} = \frac{(.6)(1-.2)}{(.6)(1-.2)+.4(1-.5)} = \frac{.48}{.68} = \frac{12}{17} = .7059$$

2. (5 pts) Find b if the pdf for the continuous random variable X is

$$f_X(x) = b + x, \qquad 0 \le x \le b$$

$$\int_0^b (b+x)\,dx = \left[bx + \frac{x^2}{2}\right]_0^b = b^2 + \frac{b^2}{2} = \frac{3}{2}b^2 = 1 \quad \Rightarrow \quad b = \sqrt{\frac{2}{3}}$$

3. (4 pts) Suppose the cdf of a discrete random variable X is

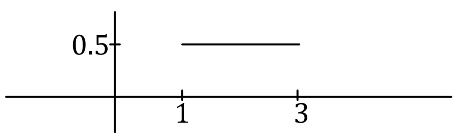
$$F_X(k) = \frac{k^2}{16}, \qquad k = 1,2,3,4$$

Find the pdf of X, fill in the following table, with numbers

k	1	2	3	4
$p_X(k)$	1/16	3/16	5/16	7/16

$$p_X(1) = F_X(1) = 1/16, \qquad p_X(2) = F_X(2) - F_X(1) = 3/16$$
$$p_X(3) = F_X(3) - F_X(2) = 5/16, \qquad p_X(4) = F_X(4) - F_X(3) = 7/16$$

4. (5 pts) Let X be a random variable with the following pdf. Find the cdf of X.



For $1 \le x \le 3$,

$$F_X(x) = \int_1^x \frac{1}{2}dt = \left[\frac{t}{2}\right]_1^x = \frac{x}{2} - \frac{1}{2}$$

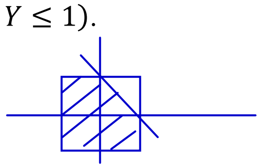
$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{2} - \frac{1}{2}, & 1 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

5. (4 pts) Let X and Y be random variable with the same pdf. Given that $E(X) = E(XY) = 2$, show that X and Y are negatively correlated.

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 2 - 2^2 = -2$$

The sign of $Cov(X,Y)$ shows that X and Y are negatively correlated.

6. (5 pts) Given that X and Y are independent continuous random variables with uniform distribution in $[-1,1]$, find $P(X+Y \le 1)$.

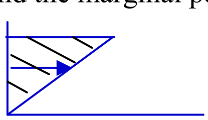


$P(X + Y \le 1) = 7/8$

7. (8 pts) Suppose the joint pdf of the random variables X and Y are

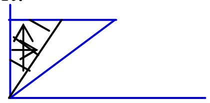
$$f_{X,Y}(x,y) = \frac{1}{2}xy, \qquad 0 < x < y < 2$$

a) (4 pts) Find the marginal pdf of Y.



$$f_Y(y) = \int_0^y \frac{1}{2}xy\,dx = \left[\frac{x^2 y}{4}\right]_0^y = \frac{y^3}{4}, \qquad 0 < y < 2$$

b) (4 pts) Set up an integral, complete with integration limits, to find $P(Y > 2X)$, but do **NOT** evaluate it.



$$P(Y > 2X) = \int_0^1 \int_{2x}^2 \frac{1}{2}xy\,dy\,dx \quad or \quad \int_0^2 \int_0^{y/2} \frac{1}{2}xy\,dx\,dy$$

8. (6 pts) The length of a large batch of 1-inch nails produced by a manufacturing process has a normal distribution with mean $\mu = 1$ and $\sigma = 0.04$. If you pick 10 nails, what is the probability that at least 8 of them have length between 0.95 and 1.05?

Probability one nail has length between 0.95 and 1.05 is

$$p = Normcdf(0.95, 1.05, 1, 0.04) = .7887$$

Let X be the number of nails with length between 0.95 and 1.05

$P(X = 8) + P(X = 9) + P(X = 10)$

$$= \binom{10}{8}(.7887^8)(.2113^2) + \binom{10}{9}(.7887^9).2113 + \binom{10}{10}(.7887^{10}) = .6435$$

Or $P(X = 8) + P(X = 9) + P(X = 10) = binomcdf(10, 0.7887, 8, 10)$

9. (6 pts) A portable system runs on a battery of which the lifetime is a random variable with mean 10.5 hours and standard deviation 2 hour. Forty batteries are installed in the system so that as soon as the active battery dies, the next battery takes over. Use the Central Limit Theorem (normal approximation) to find the probability that the system runs for at least 400 hours.

Let X be the total lifetime of the 40 batteries.

$$X \sim Normal(420, 160)$$

$$P(X \geq 400) = normcdf(400, \infty, 420, \sqrt{160}) = 0.943$$

10. (8 pts) X is a discrete random variable with the pdf

$$p_X(k;\theta) = \frac{\theta^{2k}}{k!}e^{-\theta^2} \qquad k = 0, 1, 2, \ldots$$

where θ is an unknown parameter. Find the maximum likelihood estimate (**MLE**) for θ based on a random sample of size n: $k_1, \ldots, k_n$.

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta^{2k_i}}{k_i!}e^{-\theta^2} \quad \Rightarrow \quad \ln L = \sum_{i=1}^{n} 2k_i \ln\theta - n\theta^2 - \sum_{i=1}^{n} \ln k_i!$$

$$\frac{d\ln L}{d\theta} = \frac{\sum_{i=1}^{n} 2k_i}{\theta} - 2n\theta = 0 \quad \Rightarrow \quad \theta = \sqrt{\frac{\sum_{i=1}^{n} k_i}{n}}$$

11. (4 pts) The Probability and Statistics quiz 1 scores are reported on a scale from 55 to 100. The distribution of quiz scores are approximately Normal with mean $\mu = 80$ and standard deviation $\sigma = 6$. What quiz scores make up the top 15% of all scores?

$$P(X \geq a) = 0.15 \quad \Rightarrow \quad a = invNorm(.85, 80, 6) = 86.22$$

12. (6 pts) The hemoglobin level for adult males is supposed to be about 13.8 g/dL. A sample of 20 male cardiac patients has a mean of 12.35 with $s = 2.8$. Test at the 5% level to see if the hemoglobin level for male cardiac patients is lower.

a) (2 pts) Write the null and alternate hypotheses.

$$H_0: \mu = 13.8$$
$$H_1: \mu < 13.8$$

b) (4 pts) Use **P-value** to decide if you accept or reject the null hypothesis and explain what that means.

T-test: P-value $= 0.01595 < 0.05 = \alpha$

Reject H_0: The hemoglobin level for male cardiac patients is likely to be lower.

13. (7 pts) A magazine is testing two types of computers to see if they crash at different rates. A sample of 200 of computer A finds that 12% crash when run through a rigorous set of programs. A sample of 200 of computer B finds that 8% crash when run through the same set of programs. Test at the 5% level of significance.

a) (2 pts) Write the null and alternate hypotheses.

$$H_0: p_1 = p_2 \quad \text{or} \quad p_X = p_Y$$
$$H_1: p_1 \neq p_2 \quad \text{or} \quad p_X \neq p_Y$$

b) (5 pts) Use **test statistic** to decide if you accept or reject the null hypothesis and explain what that means here.

2PropZTest: $z = 1.333 < z_{0.025} = 1.96$

Fail to reject H_0; they probably crash at the same rate.

14. (5 pts) A group wants to know what percent of people can run a mile without stopping. How large a sample will they need if they want a margin of error less than 2% at the 90% confidence level?

$$n = \frac{{z_{\alpha/2}}^2}{4d^2} = \frac{1.645^2}{4(0.02)^2} = 1691.27 \quad \Rightarrow \quad n = 1692$$

15. (8 pts) In a study of life expectancy, a random sample of 9 men lived an average of 68.2 years with a sample standard deviation of 6.2 years, and a random sample of 11 women lived an average of 70.5 years with a sample standard deviation of 7.7 years.

a) (6 pts) Test at the 0.05 significance level to see if the mean life expectancies are the same for men and women. Assume two populations have the equal variance. You should give your null and alternate hypotheses, then decide if you accept or reject the null hypothesis, by using the **test statistic**.

$$H_0: \mu_1 = \mu_2 \quad \text{or} \quad \mu_X = \mu_Y$$
$$H_1: \mu_1 \neq \mu_2 \quad \text{or} \quad \mu_X \neq \mu_Y$$

2SampleTTest (pooled): $t = -0.7235$, $df = 18$

$$|t| = 0.7235 < 2.101 = t_{0.025,18}, \quad or \quad t > -t_{0.025,18}$$

Fail to reject H_0: The mean life expectancies are probably the same for men and women.

b) (2 pts) Construct a 90% confidence interval for the difference between the two means. Do **not** assume that the two populations have the same variance.

2SampleTInt (pooled: NO)

$$(-7.69, 3.09)$$

16. (12 pts) Suppose the sales of a product per day is a normal variable with mean 400 and standard deviation 100. The company launches an ad campaign. They will look at a sample of 50 days to see if the ad campaign increases sales. They test at the 5% level of significance and use

$$H_0: \mu = 400$$
$$H_1: \mu > 400$$

a) (5 pts) Find the critical value and use it to get a rejection test.

$$x^* = invNorm\left(0.95, 400, \frac{100}{\sqrt{50}}\right) = 423.262$$

Reject H_0 if $\bar{x} > x^*$.

b) (5 pts) Find the probability of committing a Type II error if the real mean is 450.

$$\beta = P(\bar{X} > 423.262 | \mu = 450) = normcdf\left(-\infty, 423.3, 450, \frac{100}{\sqrt{50}}\right) = 0.0295$$

c) (2 pts) Give the power of the test if the real mean is 450.

$$Power = 1 - \beta = 1 - 0.0295 = 0.9705$$