1. (16 pts) Suppose X is a discrete random variable with the pdf

$$P(X=k)=p_X(k)=\theta(1-\theta)^k \quad \text{for } k=0,1,2,3,\ldots$$

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

$$L(\theta)=\prod_{i=1}^{n}\theta(1-\theta)^{k_i}$$

$$\ln L(\theta)=\sum_{i=1}^{n}\ln\theta+k_i\ln(1-\theta)$$

$$\frac{d}{d\theta}\ln L(\theta)=\sum_{i=1}^{n}\frac{1}{\theta}-\sum_{i=1}^{n}\frac{k_i}{1-\theta}=\frac{n}{\theta}-\frac{\sum_{i=1}^{n}k_i}{1-\theta}$$

$$\frac{d}{d\theta}\ln L(\theta)=0 \quad\Rightarrow\quad \frac{n}{\theta}=\frac{\sum_{i=1}^{n}k_i}{1-\theta} \quad\Rightarrow\quad n(1-\theta)=\theta\sum_{i=1}^{n}k_i \quad\Rightarrow\quad \theta\left(n+\sum_{i=1}^{n}k_i\right)=n$$

$$\theta=\frac{n}{n+\sum_{i=1}^{n}k_i}=\frac{1}{1+\bar{k}}$$

2. (12 pts) Suppose $E(X)=\theta/2$, $E(X)=\theta$, and $Var(X)=Var(Y)=\theta$. Consider the estimator

$$\hat{\theta}=2cX+(1-c)Y \qquad \text{with } 0<c<1$$

a) (5 pts) Show that $\hat{\theta}$ is unbiased for all values of c.

$$E(X)=\frac{\theta}{2}, \qquad E(Y)=\theta$$

$$E(\hat{\theta})=2cE(X)+(1-c)E(Y)=2c\left(\frac{\theta}{2}\right)+(1-c)\theta=\theta$$

b) (7 pts) Find the value of c which minimizes the variance of $\hat{\theta}$.

$$Var(\hat{\theta})=(2c)^2Var(X)+(1-c)^2Var(Y)=4c^2\theta+(1-c)^2\theta$$

$$\frac{d}{dc}Var(\hat{\theta})=8c\theta-2(1-c)\theta=0 \quad\Rightarrow\quad c=\frac{1}{5}$$

3. (12 pts) A roulette wheel is supposed to land on red 9 out of 19 times (or 47.37%). Fred found that the roulette wheel he recently bought landed on red 305 times in 600 spins.

a) (5 pts) Construct a 95% confidence interval for the proportion of landing on red for Fred's roulette wheel. Show the formula or the calculator function you use.

1-PropZInt

$$(0.4683, 0.5483)$$

b) (2 pts) Assess if the wheel is working properly based on your result in part (a).

Yes, 47.37% is within the interval.

c) (5 pts) If he wants the margin of error to be 1%, how many spins should he make?

$$n=\frac{{z_{\alpha/2}}^2}{4d^2}=\frac{1.96^2}{4(0.01)^2}=9604$$

4. (15 pts) Suppose you collected a sample of 100 data points from a normal distribution with unknown mean μ and known $\sigma=2.4$, and calculated a 95% Z confidence interval.

a) (5 pts) What is the width of the confidence interval?

$$2\left(\frac{z_{.025}\sigma}{\sqrt{n}}\right)=\frac{2(1.96)(2.4)}{\sqrt{100}}=.9408$$

b) (5 pts) If two independent samples of size 100 are collected and two 95% intervals are constructed, what is the probability that only one of intervals contain the true mean?

$$2(0.95)(0.05)=.095$$

c) (5 pts) Suppose we want to keep the confidence level at 95% but halve the width, what is the sample size we should have?

$$\frac{z_{.025}\sigma}{\sqrt{n}}=\frac{1}{2}\frac{z_{.025}\sigma}{\sqrt{100}} \quad\Rightarrow\quad n=\left(2\sqrt{100}\right)^2=400$$

5. (13 pts) A researcher claims that 25% of the U.S. population has circulation problems, but a rival researcher claims that there may be evidence in her study that the percentage should be lower, according to a random sample of 189 persons, where 38 are found to have circulation problems.

a) (8 pts) State your null and alternative hypothesis. What do you conclude at the $\alpha=0.05$ significance level about the claim of the rival research? Use the P-value.

$$H_0: p = .25 \qquad H_1: p < .25$$
$$P - value = 0.06 > \alpha = 0.05$$

At $\alpha = 0.05$, the null hypothesis is not rejected. The rival research is probably wrong.

b) (5 pts) For what values of α will the null hypothesis be rejected?

$$\alpha > P\text{-value} = 0.06$$

6. (12 pts) Suppose we test an unknown proportion p with $H_0: p = 0.3$ against $H_1: p > 0.3$ at $\alpha = 0.1$. The sample size is 500.

a) (6 pts) If the number of success is 161, do we reject H_0? Use the test statistic.

1-PropZ-Test (calculator): $z = 1.0735 < z_{0.1} = invNorm(0.9) = 1.281 \Rightarrow Fail\ to\ reject\ H_0$

b) (6 pts) What is the smallest number of successes that will cause H_0 to be rejected?

$$\frac{\bar{k} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} > z_{0.1} = 1.28155 \Rightarrow \bar{k} = \frac{x}{n} > 0.3 + 1.28155\sqrt{\frac{0.3(0.7)}{500}} = .3263$$

$$\Rightarrow x > 500(.3263) = 163.1 \Rightarrow smallest\ x = 164$$

7. (6 pts) We test the null hypothesis $H_0 : \mu = 25$ against $H_1 : \mu \neq 25$ based on a random sample of size $n = 30$ drawn from a normal distribution with $\sigma = 8$. The decision rule is to reject H_0 when $\bar{x} < 23$ or $\bar{x} > 27$, where $\bar{x}$ is the sample mean. Find the significance level α of the test.

$$\alpha = P(\bar{X} < 23 \text{ or } \bar{X} > 27 | \mu = 25) = 2 \times Normcdf\left(27, \infty, 25, 8/\sqrt{30}\right) = 0.1709$$

8. (14 pts) We test the null hypothesis $H_0 : \mu = 105$ against $H_1 : \mu > 105$ based on a random sample of size $n = 55$ drawn from a normal distribution with $\sigma = 18$ at the significance level $\alpha = 0.01$.

a) (5 pts) Find the critical value.

$$x^* = invNorm\left(0.99, 105, \frac{18}{\sqrt{55}}\right) = 110.646$$

b) (9 pts) Find the probability of committing a Type II error **and** the power of the test if the true mean is 112.

$$\beta = P(\bar{X} \leq 110.646 | \mu = 112) = normcdf\left(-\infty, 110.646, 112, \frac{18}{\sqrt{55}}\right) = 0.28847$$

$$Power = 1 - \beta = 1 - 0.28847 = 0.71153$$