UNIVERSITY OF MASSACHUSETTS
Department of Mathematics and Statistics
Advances Statistics Exam - Version 1
Wednesday, August 31, 2022

Show all work in your solution to each problem. 75 are required to pass at the Ph.D. level.

1. Suppose $X \sim Beta(a, 1)$ with pdf $f(x) = ax^{a-1}, a > 0, 0 \leq x \leq 1$.
   (a) Let $Y = 1/X$. Find the pdf of Y
   (b) Let k be a positive integer. Find $E(Y^k)$.
   (c) Y has a Pareto distribution with shape parameter a. The 80-20 rule, also known as the Pareto Principle, says that 20% of the population has 80% of the wealth. How does the Pareto Principle (approximately) relate to the Pareto distribution?

2. A response y_i is observed twice at each of at three different time points (the six t_is are $(0, 30, 60, 0, 30, 60)$, for instance). Consider two linear models for these data:

$$\begin{aligned} \text{Model 1: } y_i &= \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \varepsilon_i, i = 1, \ldots, 6, \text{ and} \\ \text{Model 2: } y_i &= \alpha_0 1_{t_i=0} + \alpha_0 1_{t_i=30} + \alpha_0 1_{t_i=60} + e_i, i = 1, \ldots, 6. \end{aligned}$$

   (Assume the errors are iid. $1_{t_i=0}$ etc. are indicator functions.)
   (a) Call the design matrix for model 1 X and the design matrix for model 2 Z. What are X and Z?
   (b) Derive the predicted responses for model 2. (Please give a very specific form for each, not just a matrix equation.)
   (c) Suppose that MSE = 1. What is the estimated covariance matrix of the estimated αs? Are the estimates independent? Why or why not?
   (d) Are the residuals and predicted responses different between the two models? Why or why not.
   (e) What is the relationship between the estimated α_ks and the estimated β_ks, $k = 1, 2, 3$?

3. Suppose that (X, Y) are bivariate normally distributed with mean $\mu = (\mu_1, \mu_2)$ and covariance $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$.
   (a) What is the marginal distribution of Y? You do not need to prove or derive this, just state the result based on knowledge of the normal distribution.
   (b) Use part (a) to show that the conditional distribution of X given $Y = y$ is $N\left(\mu_1 + \frac{\Sigma_{12}}{\Sigma_{22}}(y - \mu_2), \frac{|\Sigma|}{\Sigma_{22}}\right)$. Provide the full derivation.

(c) Express $\frac{\Sigma_{12}}{\Sigma_{22}}$ and $\frac{|\Sigma|}{\Sigma_{22}}$ in terms of the correlation coefficient $\rho = \text{Cor}(X, Y)$ and the marginal variances $\sigma_1^2 = \text{Var}(X)$ and $\sigma_2^2 = \text{Var}(Y)$.

(d) Find constants a and b (depending on μ, Σ, and y) such that the conditional distribution of $aX + b$ given $Y = y$ is $N(0, 1)$.

4. Let X_1 and X_2 be independent Unif$(\theta, \theta+1)$ observations, where θ is unknown. We wish to test the null hypothesis $H_0 : \theta = 0$ versus the **one-sided alternative** $H_A : \theta > 0$. Consider the following two tests:

$$\text{Test 1: Reject } H_0 \text{ if } X_1 > 0.95.$$
$$\text{Test 2: Reject } H_0 \text{ if } X_1 + X_2 > C.$$

(a) Find the CDF of $X_1 + X_2$ under sampling from θ.

(b) Find the value of C such that Test 2 has the same size as Test 1.

(c) Find the power function of each test (for Test 2 use the value of C you found in part (b)).