Stanford Stats 116 Final Examination

Closed book and closed notes
Duration: 180 minutes

Spring 2021

Name: ____________________

Student ID Number: ____________________

By taking this exam, you agree to be bound by the Stanford Honor Code, meaning specifically in this context that you

- will not give or receive aid in examinations;
- will do your share and take an active part in seeing to it that others as well as yourself uphold the spirit and letter of the Honor Code.

Signature: ____________________

Question F.1 (10 points)**:** Define any two of the following concepts: independence, covariance, conditional independence.

Question F.2 (15 points)**:** Assume $P(A) > 0$ and $P(B) > 0$. Prove or disprove:

(a) If $P(A \mid B) > P(A)$ then $P(B \mid A) > P(B)$.

(b) If $P(A) = P(B)$ then $P(A \mid C) = P(B \mid C)$ for any event C.

(c) If $P(A \mid B) = P(B \mid A)$ then $P(A) = P(B)$.

Question F.3 (5 points): Consider a discrete uniform random variable over the set $S = \{1, 3, -1, -2, 5, 7\}$. What is $\mathbb{E}[X \mid X > 0]$?

Question F.4 (10 points): An urn contains 10 balls of which three are black and seven are white. The following game is played: At each trial a ball is selected at random, its color is noted, and it is replaced along with two additional balls of the same color. What is the probability that a black ball is selected in each of the first three trials?

Question F.5 (10 points): Let X and Y be independent and identically distributed as $\mathsf{Uni}[0, 1]$. Recall that the pdf of $\mathsf{Uni}[0, 1])$ is $f(x) = \mathbf{1}\{0 \leq x \leq 1\}$.

(a) What is the distribution of $-\log(X)$?

(b) What is the distribution of $\log(XY)$? You can recall a known result from class.

Question F.6: Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ be independent and identically distributed with $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = 0$ and $\text{Var}(X_1) = \text{Var}(Y_1) = 1$. Define $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$ to be the corresponding sample averages.

(a) (5 points) Using Chebychev's inequality, what can you say about $P(-1 \leq \overline{X}_n - \overline{Y}_n \leq 1)$?

(b) (5 points) Using the central limit theorem, approximate $P(-1 \leq \overline{X}_n - \overline{Y}_n \leq 1)$ for $n = 50$.

Question F.7 (10 points)**:** A child who doesn't want to go to school picks *headache* as reason with probability 0.7 and *stomach-ache* as reason with probability 0.3. If she picks *headache*, then she misses school for $\mathsf{Poisson}(1)$ days, while if she picks *stomach-ache*, then she misses school for $\mathsf{Poisson}(2)$ days. The teacher observes that the child has not come to school for 2 days and tries to guess the chance that she will put in *stomach-ache* as reason. Help the teacher by giving this chance. Recall if $X \sim \mathsf{Poisson}(\lambda)$, then $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$.

Question F.8: A stochastic linear dynamical system evolves according to

$$X_{n+1} = aX_n + bW_n$$

where $W_n \overset{\text{iid}}{\sim} \mathsf{N}(0,1)$ are called *innovations*, and $0 \le a < 1$ and $b \in \mathbb{R}$ are fixed constants.

(a) (5 points) If $X_n \sim \mathsf{N}(0, \sigma^2)$, what is the distribution of X_{n+1}?

(b) (5 points) Give a probability distribution P so that if X_n has distribution P, then so does X_{n+1}. (Note that they are *not* independent, but they have the same distribution.)

(c) (5 points) Suppose $X_0 = 0$. Let P_n be the probability distribution of X_n. Give the limiting distribution $\lim_{n\to\infty} P_n$.

Question F.9 (15 points): Consider the bivariate function

$$f(x, y) = K \cdot (x + y) \cdot \mathbf{1}\{0 \leq x \leq 1\} \, \mathbf{1}\{0 \leq y \leq 1\}.$$

(a) Find the constant K such that $f(x, y)$ is a joint probability density function.

(b) If (X, Y) has density $f(x, y)$ find the density $f_X(x)$ of X.

(c) Find $f_{Y|X}(y \mid x)$.

Question F.10 (10 points)**:** Recall that the pdf of $\mathsf{N}(\mu, \sigma^2)$ is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$. Suppose that $X_1, X_2, \cdots, X_n$ are independent normal random variables with $X_i \sim \mathsf{N}(0, \sigma_i^2)$ for some $\sigma_i > 0$ (that is, the standard deviation of X_i is σ_i). Let $Z = \sum_{i=1}^n X_i$.

(a) Compute the moment generating function (MGF) of Z.

(b) Using the Chernoff bound, show that for any $a > 0$,

$$P(Z \geq a) \leq e^{-\frac{a^2}{2\sum_{i=1}^n \sigma_i^2}}.$$