UNIVERSITY OF CALIFORNIA, BERKELEY

DEPARTMENT OF STATISTICS

STAT 134: Concepts of Probability

Spring 2014

Instructor: Antar Bandyopadhyay

Solution to the Final Examination

1. State whether the following statements are *true* or *false*. Write brief reasons supporting your answers in the space provided. **For each correct guess you will get +1 points. If your guess is correct and your reasoning is also correct then you will get an additional +4 points**. $[(1+4) \times 4 = 20]$

(a) For any two events A and B we must have $\mathbf{P}(A \cap B) \geq \mathbf{P}(A) - \mathbf{P}(B^c)$.

Answer: TRUE

Argument in favor of the answer:

$$\mathbf{P}(A) - \mathbf{P}(B^c) = \mathbf{P}(A) + \mathbf{P}(B) - 1 \leq \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cup B) = \mathbf{P}(A \cap B).$$

(b) Suppose X and Y are two independent random variables with Exponential (1) distribution. Then $|X - Y|$ does not have *memorilessness* property.

Answer: FALSE

An argument in favor of the answer: Since X and Y are *i.i.d.* random variables with Exponential (1) distribution then the distribution of $X - Y$ is *Double Exponential* with parameter 1, that is, the density of $X - Y$ is given by

$$f_{X-Y}(t) = \frac{1}{2} e^{-|t|}, \ \ t \in \mathbb{R}.$$

Thus $Z = |X - Y|$ is again Exponential (1) which has *memorylessness property*.

(c) Suppose $X \sim \text{Normal}(0, 1)$. Then for $n, m \in \{1, 2, 3, \ldots\}$ the random variables X^n and X^m are *uncorrelated*, that is correlation is zero, if and only if, $n + m$ is an odd number.

Answer: TRUE

An argument in favor of the answer: Observe that if $X \sim \text{Normal}(0,1)$ then

$$\mathbf{E}\left[X^k\right] = \begin{cases} 1 \times 3 \times \cdots \times (k-1) & \text{if } k \text{ is even;} \\ 0 & \text{if } k \text{ is odd.} \end{cases}$$

Thus $\text{Cov}(X^n, X^m) = 0$, if and only if, $m+n$ is an odd number.

(d) If X and Y are two non-negative random variables such that $\mathbf{E}\left[Y \,\middle|\, X\right] \geq \mathbf{E}[Y]$ then X and Y have non-negative correlation.

Answer: TRUE

An argument in favor of the answer: As X is non-negative so it follows from what is given that

$$X\,\mathbf{E}\left[Y \,\middle|\, X\right] \geq X\,\mathbf{E}[Y].$$

Taking expectations on both side we can conclude that

$$\mathbf{E}[XY] \geq \mathbf{E}[X]\,\mathbf{E}[Y].$$

2. A point (X,Y) is uniformly selected from the following region on the plane

$$\left\{(x,y) \,\middle|\, \sqrt{|x|} + \sqrt{|y|} \leq 1\right\}.$$

(a) What is the marginal distributions of X? [5]

The joint density of (X,Y) is given by

$$f_{(X,Y)}(x,y) = \frac{3}{2}\mathbf{1}\left(\sqrt{|x|} + \sqrt{|y|} \leq 1\right).$$

Because the area of the given region is $4 \times \int_0^1 (1-\sqrt{x})^2\, dx = 4 \times \frac{1}{6} = \frac{2}{3}$. Thus the marginal density of the random variable X is given by

$$f_X(x) = 3\left(1 + x - 2\sqrt{|x|}\right)\mathbf{1}(|x| \leq 1).$$

(b) Find the conditional distribution of Y given the event $\left[X = \frac{1}{4}\right]$. [5]

The conditional distribution of Y given the event $\left[X = \frac{1}{4}\right]$ is uniform on the interval $\left(-\frac{1}{4}, \frac{1}{4}\right)$.

3. Suppose a random number U is generated out of $\text{Uniform}(0,1)$ distribution. Once U is observed, say $U = u$, then we toss a coin with probability of head u, repeatedly till we get a head. Let X be the number of tosses before the first head.

(a) Find the marginal distribution of X. [5]

Notice that $X \Big| U = u \sim \text{Geometric}\,(u)$. Thus the random variable X takes values in the set $\{1, 2, 3, \ldots, \}$. Also for $k \geq 1$,

$$\begin{aligned}
\mathbf{P}\,(X = k) &= \int_0^1 \mathbf{P}\left(X = k \Big| U = u\right) du \\
&= \int_0^1 u\,(1-u)^{k-1}\, du \\
&= \text{Beta}\,(2, k) \\
&= \frac{1}{k\,(k+1)}.
\end{aligned}$$

(b) Find the conditional distribution of U given the event $[X = 10]$. [5]

Observe that

$$\mathbf{P}\left(U \leq u \Big| X = 10\right) = 10 \times 11 \int_0^u t\,(1-t)^9\, dt.$$

So the conditional distribution of U given $X = 10$ is $\text{Beta}\,(2, 10)$.

4. The Shuttack Theater in Berkeley is doing a retrospective on the Harry Potter movies by selling tickets only from the counter. On a Saturday evening Potter fans arrive at the counter at a rate 4 per 10 minutes. It is known that 75% of the Potter fans in Berkeley are teenagers. Let X be the total number of teenagers who arrive in the Shuttack Theater during the 3 hours of the evening. Find $\mathbf{E}\,[X]$ and $\mathbf{Var}\,(X)$.

[5 + 5 = 10]

Let X_i be the indicator/Bernoulli variable indicating the event that the i^{th} customer is a teenager. Note that $\mathbf{P}\,(X_i = 1) = \mathbf{E}\,(X_i) = \frac{3}{4}$. Let N be the total number of customers who arrive in the Shuttack Theater during the 3 hours of the evening. Then first of all $N \sim \text{Poisson}\,(\lambda)$ where $\lambda = 3 \times 60 \times \frac{4}{10} = 72$. Further from definition

$$X = \sum_{i=1}^{N} X_i.$$

It is done in class that $\mathbf{E}\,[X] = 72 \times \frac{3}{4} = 54$ and also $\mathbf{Var}\,(X) = 54$. In fact $X \sim \text{Poisson}\,(54)$.

5. Suppose X and Y are independent and identically distributed $\text{Normal}\,(0, 1)$ random variables. Find the *probability density function* and the *cumulative distribution function* of the random variable

$$Z = \frac{X}{|Y|}.$$

[5 + 5 = 10]

Notice that as $Y \sim \text{Normal}\,(0, 1)$ so $|Y|$ and $\text{sign}\,(Y)$ are independent. We know that $\frac{X}{Y} \sim \text{Cauchy}\,(0, 1)$ which is a *symmetric* distribution. Thus $Z = \frac{X}{|Y|}$ is also $\text{Cauchy}\,(0, 1)$ random variable.

6. An urn contains 50 green balls, 25 red balls and 25 yellow balls. Balls are being selected out of the urn **with replacement**. The sampling is done till 10 balls of color red are obtained. Let X be the total number of sample and Y be the number of green balls in the sample.

   (a) Find $\mathbf{E}\left[e^X\right]$. [5]

   First note that $X \sim$ Negative-Binomial $\left(10, \frac{1}{4}\right)$. Thus $X = Y_1 + Y_2 + \cdots + Y_{10}$ where $(Y_i)_{i=1}^{10}$ are *i.i.d.* Geometric $\left(\frac{1}{4}\right)$. So

   $$\mathbf{E}\left[e^X\right] = \left(\mathbf{E}\left[e^{Y_1}\right]\right)^{10}.$$

   But $\mathbf{E}\left[e^{Y_1}\right] = \frac{1}{4}\sum_{k=1}^{\infty}\left(\frac{3\,e}{4}\right)^k = \infty$, as $3\,e > 6 > 4$. So $\mathbf{E}\left[e^X\right] = \infty$.

   (b) What is the conditional distribution of Y given $X = 50$. [5]

   From direct computation it follows that the conditional distribution of Y given $X = 50$ is Binomial $\left(40, \frac{2}{3}\right)$. Note this problem is exactly similar to the Problem # 4(b) of the Practice Midterm.

7. Balls are being drawn at random without replacement from a box containing n balls with numbers $1, 2, \ldots, n$ written on them. Let X_i be the number on the ball drawn at the i^{th} stage.

   (a) Find the Conditional distribution of $(X_1, X_2, \ldots, X_{n-1})$ given the event $[X_n = n]$. [2]

   The conditional distribution of $(X_1, X_2, \ldots, X_{n-1})$ given the event $[X_n = n]$ is uniform on the set of permutations of the numbers $\{1, 2, \ldots, n-1\}$.

   (b) Find $\text{Corr}\,(X_1, X_2)$. [3]

   Note $\mathbf{E}\,[X_1] = \frac{n+1}{2} = \mathbf{E}\,[X_2]$, $\mathbf{Var}\,(X_1) = \frac{n^2-1}{12} = \mathbf{Var}\,(X_2)$. Also observe that $X_1 + X_2 + \cdots + X_n = \frac{n(n+1)}{2}$. Thus $\mathbf{Var}\,(X_1 + X_2 + \cdots + X_n) = 0$. But from the variance sum formula and exchangeability of $(X_1, X_2, \ldots, X_n)$ we then get

   $$n\mathbf{Var}\,(X_1) + n\,(n-1)\,\text{Cov}\,(X_1, X_2) = 0.$$

   This gives $\text{Cov}\,(X_1, X_2) = -\frac{n+1}{12}$. So $\text{Corr}\,(X_1, X_2) = -\frac{1}{n-1}$.
   Note that the covariance of X_1 and X_2 can also be computed directly.

   (c) Find $\mathbf{Var}\,(X_1 + X_2 + \cdots + X_{n-1})$. [5]

   Observe that $X_1 + X_2 + \cdots + X_{n-1} = \frac{n(n-1)}{2} - X_n$ and hence $\mathbf{Var}\,(X_1 + X_2 + \cdots + X_{n-1}) = \mathbf{Var}\,(X_n) = \mathbf{Var}\,(X_1) = \frac{n^2-1}{12}$.

8. Suppose (X, Y) is a pair of random variables with joint density given by

   $$f\,(x, y) = \begin{cases} c\,x^{\alpha-1}\,y^{\beta-1}\,(1-x-y)^{\gamma-1} & \text{if } 0 < x,\ y,\ x+y < 1; \\ 0 & \text{otherwise,} \end{cases}$$

   where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$.

(a) Find c. [3]

By direct integration we will get $c = \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}$.

(b) Find the marginal distribution of X. [3]

The marginal distribution of X is Beta $(\alpha, \beta+\gamma)$.

(c) Find $\mathbf{E}\left[Y \,\middle|\, X = x\right]$ where $0 < x < 1$. [4]

Observe that the conditional distribution of Y given $[X = x]$ is $(1-x)$ Beta (β, γ), when $0 < x < 1$. Thus

$$\mathbf{E}\left[Y \,\middle|\, X = x\right] = (1-x)\,\frac{\beta}{\beta+\gamma}.$$

9. A chestnut drawer has two drawers. 10 different pairs of socks are randomly placed in the two drawers.

(a) Let N be the number of complete pairs of socks in the first drawer. Find the distribution of N. [4]

Observe that each pair of socks can either be placed in the first drawer, or in the second drawer, or they can be split in the two drawers with probabilities $\frac{1}{4}$, $\frac{1}{4}$ and $\frac{1}{2}$ respectively. Thus from definition $N \sim \text{Binomial}\left(10, \frac{1}{4}\right)$.

(b) What is the probability that each drawer has at least one complete pair of socks? [6]

Let M be the total number of complete pairs in the second drawer. Note that M is identically distributed as N, but they are not independent. In fact, $(N, M, 10 - N - M)$ is a Multinomial $\left(10; \frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$. Now

$$\begin{aligned}
\mathbf{P}\left(\text{Each drawer has at least one complete pair of socks }\right) &= \mathbf{P}\left(N \geq 1, M \geq 1\right) \\
&= 1 - \mathbf{P}\left(N = 0 \text{ or } M = 0\right) \\
&= 1 - 2\mathbf{P}\left(N = 0\right) + \mathbf{P}\left(N = M = 0\right) \\
&= 1 - 2 \times \left(\frac{3}{4}\right)^{10} + \left(\frac{1}{2}\right)^{10} \\
&\approx 0.8883495.
\end{aligned}$$