Stat 510: Final Exam.

Moulinath Banerjee

December 21, 2015

Announcement: The exam carries 85 points but the maximum possible score is 75 points. Extra sheets must be stapled at the back. Show ALL your work. No calculators allowed. Internet connections on electronic devices being used for exam purposes MUST be turned off.

- (1) (a) Let $X = R\cos(\Theta)$ and $Y = R\sin(\Theta)$ where $R > 0$ is a random variable independent of the random variable Θ which follows a uniform distribution on $(0, 2\pi)$. Show that Y/X follows the standard Cauchy distribution.

  (b) If $X \sim \text{Uniform}(0,1)$ and conditional on $X = x$, Y is generated from $N(x, x^2)$, show that Y/X is independent of X. (10 + 10 = 20 points)

1 (a) $\frac{Y}{X} = \tan\Theta$. where $\Theta \sim \text{Uniform}(0, 2\pi)$.

$\Theta \longmapsto g(\Theta)$ is a differentiable transformation. To find the density of $\tan\Theta$ look at Exercise 2.6 on Page 22 of ~~the~~ 510NOTES - Fall2020 file on Canvas.

(b) $Y \mid X = x \sim N(x, x^2)$

so $\frac{Y}{x} \mid X = x \sim N(1, 1)$.

But the distribution of $\frac{Y}{X} \mid X = x$ is identical to ~~that~~ of $\frac{Y}{x} \mid X = x$.

So: $\frac{Y}{X} \mid X = x \sim N(1, 1)$ and this conditional distribution does not depend on x, the value of X

$$\boxed{\text{Shows } \frac{Y}{X} \text{ is independent of } X}$$

- (2) (a) Let (X, Y) follow the Bivariate Normal Distribution with parameters $\mu_x = 0, \mu_y = 0, \sigma_x = 1 = \sigma_y$ and $-1 < \rho < 1$. Show that $Y = \rho X + \epsilon$ for a random variable ϵ that is independent of X. What is the distribution of ϵ?

  (b) If $X \sim N(0, 1)$ and $Y \mid X = x$ is distributed like a $N(\rho X, 1 - \rho^2)$ random variable, what is the joint distribution of (X, Y)? (12 + 8 = 20 points)

These were all discussed in the Notes on Bivariate Normal Distributions.

2(a). Define: $U = X$

$$\mathcal{E} = Y - \rho X.$$

Consider the transformation:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \longmapsto \begin{pmatrix} U \\ \mathcal{E} \end{pmatrix}.$$

This is an invertible linear transformation.

So apply the change of variable theorem to find the joint density of $(U, \mathcal{E})$.

The joint density will factor into a function of U and one of $\mathcal{E}$, showing that U and $\mathcal{E}$ are independent. Since $U = X$, this shows that $\mathcal{E} \equiv Y - \rho X$ is independent of X.

2(b) Since the marginal of X and the conditional of Y given X are given, write down the joint density $f_{X,Y}(x,y) = f_X(x) f_{Y|X=x}(y|x)$.

By the algebra done in class (check Scribbles 22 and 23) it follows ~~that~~ $(X, Y) \sim BVN(0, 0, 1, 1, \rho)$

- (3) (a) Let X_1, X_2 be i.i.d. Uniform$(0, \theta)$ where $\theta > 0$ is an unknown parameter. Thus the common density is $f(x) = (1/\theta)\, 1(x \leq \theta)$. Let U and V denote the minimum and maximum of these two observations. It is intuitively clear that in the presence of V, U is un-informative about θ, since if you know V you know that $\theta \geq V$ and the smaller quantity U is irrelevant in the presence of V. Formalize this intuition by first showing that the joint density of (U, V) is $f(u, v) = (2/\theta^2)\, 1(0 < u < v < \theta)$ and then using this to deduce that the conditional distribution of U given V does not depend on θ.
  (b) Let $X_1, X_2, \ldots,$ be a sequence of i.i.d Exponential random variables with mean β. Let N be a Geometric random variable assuming values 1,2,... with parameter p and independent of the sequence of X_i's. Show that the distribution of $S_N = X_1 + X_2 + \ldots + X_N$ is also exponential and determine its parameter. (MGF's can be useful). (10 + 10 = 20 points)

3(a) $P(U > u, V \leq v)$ $\quad 0 < u < v < \theta$

$$= P(X_1 > u, X_2 > u, X_1 \leq v, X_2 \leq v)$$

$$= P(X_1 \in (u, v], X_2 \in (u, v])$$

$$= P(X_1 \in (u, v])\, P(X_2 \in (u, v])$$

$$= \frac{(v-u)^2}{\theta^2}$$

$$P(V \leq v) = P(U > u, V \leq v) + P(U \leq u, V \leq v)$$

so $P(U \leq u, V \leq v) = P(V \leq v) - P(U > u, V \leq v)$

$$= \frac{v^2}{\theta^2} - \frac{(v-u)^2}{\theta^2}$$

$$f(u,v) = \frac{\partial^2}{\partial u \partial v} P(U \leq u, V \leq v)$$

$$= \frac{\partial}{\partial u}\left[- \frac{2(v-u)}{\theta^2} \right] = \frac{2}{\theta^2}, \text{ for } 0 < u < v < \theta.$$

Conditional of U given $V=v$, for $v<\theta$

$$f_{U|V}(u|v) = \frac{f_{U,V}(u,v)}{f_V(v)} \mathbb{1}(0<u<v)$$

$$= \frac{\frac{2}{\theta^2}\mathbb{1}(0<u<v)}{\frac{2v}{\theta^2}}$$

$$= \frac{1}{v}\mathbb{1}(0<u<v).$$

This shows ~~that~~ $U|V=v$ follows Uniform$(0,v)$.

2(b). $\mathbb{E}[e^{tS_N}]$

$$= \mathbb{E}[\mathbb{E}[e^{tS_N}|N]]$$

$$= \sum_{n=1}^{\infty} P(N=n)\,\mathbb{E}[e^{tS_N}|N=n]$$

$$= \sum_{n=1}^{\infty} P(N=n)\,\mathbb{E}[e^{tS_n}|N=n]$$

$$= \sum_{n=1}^{\infty} P(N=n)\,\mathbb{E}[e^{tS_n}] \quad (*)$$

$$S_n = X_1 + \cdots + X_n$$

$$\mathbb{E}[e^{tS_n}] = (\mathbb{E}[e^{tX_1}])^n$$

$X_1 \sim \text{Exp}(\beta)$

$$f_{X_1}(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \mathbb{1}(x>0)$$

$$M_{X_1}(t) = \frac{1}{1-\beta t} \quad \text{for } t < \frac{1}{\beta}.$$

So, from $(*)$:

$$M_{S_N}(t) = \sum_{n=1}^{\infty} p(1-p)^{n-1}\left(\frac{1}{1-\beta t}\right)^n$$

$$= \frac{1}{1-\beta t}\sum_{n=1}^{\infty} p\left[\frac{1-p}{1-\beta t}\right]^{n-1}$$

$$= \frac{p}{1-\beta t}\left[\frac{1}{1-\frac{1-p}{1-\beta t}}\right]$$

$$\times \left[\sum_{n=1}^{\infty}\left(1-\frac{1-p}{1-\beta t}\right)\left(\frac{1-p}{1-\beta t}\right)^{n-1}\right]$$

The sum in [ ] = 1

so $M_{S_N}(t) = \frac{p}{\cancel{1-\beta t}} \cdot \frac{\cancel{1-\beta t}}{\cancel{1}-\beta t-\cancel{1}+p} = \frac{1}{1-\frac{\beta}{p}t}$

so $S_N \sim Exp\left(\frac{\beta}{p}\right)$

- (4) Let $X_1, X_2, \ldots, X_n$ be i.i.d. Bernoulli(θ) random variables. Let $S_n = X_1 + X_2 + \ldots + X_n$. Let $0 \le k \le n$.
  (i) On what set $\mathcal{S}$ of n–tuples is the conditional distribution of $(X_1, X_2, \ldots, X_n)$ given $S_n = k$ supported? Show that this conditional probability mass function is the uniform distribution on the set $\mathcal{S}$.
  (ii) Show that $E(X_1 \mid S_n = k) = k/n$. Why do you think it is increasing in k?
  (iii) Show that the conditional correlation between X_1 and X_2 given $S_n = k$ does not depend on k. (Isolating the cases $k = 0, n$ from the others might help.) $(8 + 8 + 9 = 25$ points)

4 (i) $\mathcal{S} = \{(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) : \sum \varepsilon_i = k\}$

Let $(\varepsilon_1, \ldots, \varepsilon_n) \in \mathcal{S}$ and consider

$P(X_1 = \varepsilon_1, \ldots, X_n = \varepsilon_n \mid S_n = k)$

$$= \frac{P(X_1 = \varepsilon_1, \ldots, X_n = \varepsilon_n, S_n = k)}{P(S_n = k)}$$

$$= \frac{P(X_1 = \varepsilon_1, \ldots, X_n = \varepsilon_n)}{P(S_n = k)}$$

$$= \frac{\prod_{i=1}^{n} \theta^{\varepsilon_i}(1-\theta)^{1-\varepsilon_i}}{\binom{n}{k}\theta^k(1-\theta)^{n-k}} = \frac{\theta^k(1-\theta)^{n-k}}{\binom{n}{k}\theta^k(1-\theta)^{n-k}}$$

$$= \frac{1}{\binom{n}{k}}.$$

(ii) $\mathbb{E}(X_1 + \ldots + X_n \mid S_n = k) = k$

Now: $\mathbb{E}(X_1 + \ldots + X_n \mid S_n = k)$

$$= \sum_{i=1}^{n} \mathbb{E}[X_i \mid S_n = k]$$

Now, conditional on $S_n = k$, $X_1, X_2, \ldots, X_n$ are all identically distributed.

i.e $\mathbb{E}[X_i \mid S_n = k] = \mathbb{E}[X_1 \mid S_n = k]$ for all i

so: $n\,\mathbb{E}[X_1 \mid S_n = k] = k$

$$\Rightarrow \mathbb{E}[X_1 \mid S_n = k] = \frac{k}{n}.$$

<u>(iii)</u>: $\mathbb{E}[X_1 X_2 \mid S_n = k] = \mathbb{E}[X_i X_j \mid S_n = k]$ for all $i \neq j$

so: $\mathbb{E}\left[\sum_{i \neq j} X_i X_j \mid S_n = k\right] = n(n-1) \times \mathbb{E}[X_1 X_2 \mid S_n = k]$

i.e $\mathbb{E}\left[\left\{\left(\sum X_i\right)^2 - \sum X_i^2\right\} \mid S_n = k\right] = \quad ''$

i.e $\mathbb{E}\left[(S_n^2 - S_n) \mid S_n = k \right] = n(n-1) \times \mathbb{E}\left[x_1 x_2 \mid S_n = k\right]$

(using $\sum x_i^2 = \sum x_i$)

$$\Rightarrow \quad \mathbb{E}\left[x_1 x_2 \mid S_n = k \right] = \frac{k(k-1)}{n(n-1)}.$$

$$\underline{\mathrm{Var}(x_1 \mid S_n = k)} = \frac{k}{n}\left(1 - \frac{k}{n}\right)$$

since given $S_n = k$, $x_1 \sim \mathrm{Bin}\left(\frac{k}{n}\right)$.

so: $\mathrm{Cov}(x_1, x_2 \mid S_n = k)$

$$= \frac{k(k-1)}{n(n-1)} - \frac{k^2}{n^2}$$

$\mathrm{Corrn}(x_1, x_2 \mid S_n = k)$

$$= \frac{\frac{k(k-1)}{n(n-1)} - \frac{k^2}{n^2}}{\frac{k}{n} \cdot \left(1 - \frac{k}{n}\right)}$$

$$= \frac{\frac{k-1}{n-1} - \frac{k}{n}}{1 - k/n} = \frac{\frac{\cancel{nk} - n - \cancel{nk} + k}{n(n-1)}}{\frac{n-k}{n}} = -\frac{1}{n-1}$$