Bilkent University **Fall 2016-2017**

Math255 Probability and Statistics
Midterm 2
1 Dec. 2016

Problem 1. [5 pts] Let X_A, X_E, X_K be the rainfalls (in cms) in a given year to Ankara, Eskişehir, and Konya provinces, respectively. Suppose the rainfalls are individually Gaussian random variables $X_A \sim N(40, 100)$, $X_E \sim N(50, 144)$, $X_K \sim N(30, 81)$, with pairwise correlations $\rho(X_A, X_E) = \rho(X_A, X_K) = \rho(X_E, X_K) = 0.5$. Compute $\mathrm{var}(X)$ where $X = X_A + X_E + X_K$ is the total rainfall in all three provinces. (Express your answer as a real number.)

Solution. Recall that, for any two RVs X, Y, $\mathrm{var}(X+Y) = \mathrm{var}(X) + 2\,\mathrm{cov}(X,Y) + \mathrm{var}(Y)$ and $\mathrm{cov}(X,Y) = \sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}\,\rho(X,Y)$. Thus,

$$\begin{aligned}\mathrm{var}(X) &= \mathrm{var}(X_A) + \mathrm{var}(X_E) + \mathrm{var}(X_K) + 2\,\mathrm{cov}(X_A, X_E) + 2\,\mathrm{cov}(X_A, X_K) + 2\,\mathrm{cov}(X_E, X_K)\\ &= 100 + 144 + 81 + 2(\sqrt{100\cdot 144})\,0.5 + 2(\sqrt{100\cdot 81})\,0.5 + 2(\sqrt{144\cdot 81})\,0.5\\ &= 325 + 120 + 90 + 108 = 643.\end{aligned}$$

Problem 2. [5 pts] Let X be exponentially distributed with parameter $\mu > 0$,

$$f_X(x) = \mu e^{-\mu x}, \quad x \geq 0;$$

and Y be Poisson with parameter x conditional on $X = x$,

$$p_{Y|X}(k|x) = \frac{x^k e^{-x}}{k!}, \quad k = 0, 1, \ldots.$$

Evaluate $\mathbf{E}[X|Y=1]$ for $\mu = 2$. (A real number is required as an answer.)

Solution. By definition,

$$\mathbf{E}[X|Y=1] = \int x f_{X|Y}(x|1)dx = \int x\frac{f_X(x)p_{Y|X}(1|x)}{p_Y(1)}dx = \int_0^\infty x\,\frac{\mu e^{-\mu x} x e^{-x}}{p_Y(1)}dx,$$

where

$$p_Y(1) = \int f_X(x) p_{Y|X}(1|x)dx = \int_0^\infty \mu e^{-\mu x} x e^{-x} dx = \mu \int_0^\infty x e^{-(1+\mu)x}dx = \mu(1+\mu)^{-2}.$$

Thus,

$$\mathbf{E}[X|Y=1] = \int_0^\infty x\,\frac{\mu e^{-\mu x} x e^{-x}}{\mu(1+\mu)^{-2}}dx = \int_0^\infty \frac{x^2 e^{-(1+\mu)x}}{(1+\mu)^{-2}}dx = 2(1+\mu)^{-1}.$$

For $\mu = 2$, we have $\mathbf{E}[X|Y=1] = 2/3$.

Problem 3. [5 pts] Let (X, Y) have the joint distribution

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right], \quad -\infty < x, y < \infty.$$

Let $Z = 2X + Y$. Compute the CDF $F_Z(z)$ in terms of the CDF of unit normal, $\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}dv$, and the parameters μ_1, μ_2, σ_1 and σ_2.

Solution. Since $f_{X,Y}(x,y)$ has product form, we understand that X and Y are independent with $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$. Thus, from class notes, $Z = 2X + Y$ is Gaussian with mean $\mu_z = 2\mu_1 + \mu_2$ and variance $\sigma_z^2 = 4\sigma_1^2 + \sigma_2^2$. Hence,

$$F_Z(z) = P(Z \leq z) = P\left(\frac{Z-\mu_z}{\sigma_z} \leq \frac{z-\mu_z}{\sigma_z}\right) = \Phi\left(\frac{z-\mu_z}{\sigma_z}\right) = \Phi\left(\frac{z - 2\mu_1 - \mu_2}{\sqrt{4\sigma_1^2 + \sigma_2^2}}\right).$$

Problem 4. [5 pts] Let X be uniform on $[0, \frac{\pi}{2}]$, *i.e.*, $f_X(x) = \frac{2}{\pi}$, $0 \le x \le \frac{\pi}{2}$, and let $Y = \sin(X)$. Compute $f_Y(y)$ at $y = \sqrt{3}/2$. (A real number is required as an answer, with full justification.)

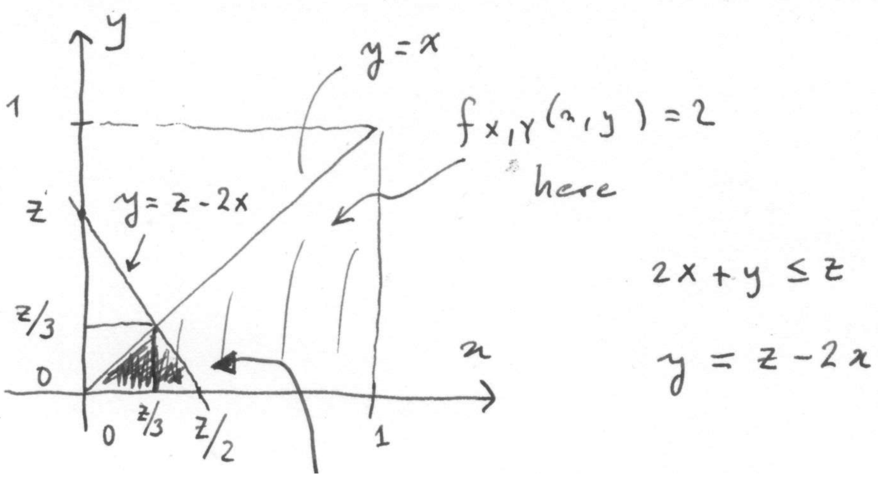
Solution. Let $y = g(x) \triangleq \sin(x)$. This function is one-to-one and increasing over $[0, \pi/2]$, which is the range of the RV X. Hence, we have

$$f_Y(y) = \frac{1}{|g'(x)|} f_X(x) = \frac{2/\pi}{\cos(x)} = \frac{2/\pi}{\sqrt{1-y^2}}, \quad 0 \le x \le \pi/2.$$

Thus, $f_Y(\sqrt{3}/2) = 4/\pi$.

Problem 5. [5 pts] Let (X, Y) be jointly distributed uniformly over the triangle in the (x, y) plane with corners at $(0,0)$, $(1,1)$, $(1,0)$. Let $Z \triangleq 2X + Y$ and compute $f_Z(z)$ for $z = 1$.

Solution. The joint density equals 2 over the region $0 \le y \le x \le 1$, and 0 elsewhere. The range of $Z = 2X + Y$ is 0 to 3. We will compute $f_Z(1)$ by first computing $F_Z(z)$ as a function of z and then computing $f_Z(1) = \mathrm{d}F(1)/\mathrm{d}z$. The figure shows how to calculate $F_Z(z)$ as a function of z.



$F_Z(z)$ equals the shaded area times the density (2) for $0 \le z \le 2$, in other words,

$$F_Z(z) = \frac{1}{2} \times \frac{z}{2} \times \frac{z}{3} \times 2 = \frac{z^2}{6}, \quad 0 \le z \le 2.$$

Differentiating, we get

$$f_Z(1) = (2z/6)\Big|_{z=1} = 1/3.$$

Problem 6. [5 pts] In 100 independent tosses of a fair die (with faces numbered 1 through 6), let X_i be the outcome (number on the up face) in the ith trial and $M = \frac{1}{100}\sum_{i=1}^{100} X_i$ be the empirical mean. Estimate the $P(3.4 < M < 3.6)$ using the central limit theorem and the table below (use linear interpolation if the number you are seeking is not in the table). A numerical answer is required with 4 digit precision. Show details on the solution page.

Solution. Observe that $\mathbf{E}[X_i] = 3.5$ and $\mathrm{var}(X_i) = \sum_{i=1}^{6} \frac{1}{6}(i - \frac{1}{6}\sum_{i=1}^{6} i)^2 = 35/12$. Thus,

$\mathbf{E}[M] = \mathbf{E}[X_i] = 3.5$ and $\mathrm{var}(M) = \mathrm{var}(X_i)/100 = 35/1200$. Now, we can use the CLT.

$$
\begin{aligned}
P(3.4 < M < 3.6) &= P\left(\frac{3.4 - 3.5}{\sqrt{35/1200}} < \frac{M - 3.5}{\sqrt{35/1200}} < \frac{3.6 - 3.5}{\sqrt{35/1200}}\right) \\
&= P\left(\frac{-0.1}{\sqrt{35/1200}} < \frac{M - 3.5}{\sqrt{35/1200}} < \frac{0.1}{\sqrt{35/1200}}\right) \\
&= P\left(-0.5855 < \frac{M - 3.5}{\sqrt{35/1200}} < 0.5855\right) \\
&\approx 2\,\Phi\left(\sqrt{\frac{12}{35}}\right) - 1 = 2\,\Phi(0.5855) - 1 = 2 \times 0.7209 - 1 = 0.4418
\end{aligned}
$$