STAT 712 hw 5

Joint and marginal distributions, conditional distributions, independence

Do problems 4.1, 4.9, 4.10, 4.11, 4.15 from CB. In addition:

1. A frog will hop across a sidewalk, beginning the dirt on one side and ending in the dirt on the other side. Let X be the number of times the frog lands on the sidewalk while hopping across. Derive the probability mass function of X assuming that the frog's hopping distances are independent and have the exponential distribution with mean $1/\lambda$ and that the sidewalk has width t.

Let $Y_1, Y_2, \dots$ be the hopping distances of the frog. First consider finding $P(X = 0)$. We have

$$P(X = 0) = P(Y_1 > t) = 1 - (1 - e^{t\lambda}) = e^{t\lambda}.$$

Now, for any $k \geq 1$, we have

$$\begin{aligned} P(X = k) &= P(\{Y_1 + \cdots + Y_{k+1} > t\} \cap \{Y_1 + \cdots + Y_k \leq t\}) \\ &= P(\{G_k + Y_{k+1} > t\} \cap \{G_k \leq t\}), \quad G_k \sim \text{Gamma}(k, 1/\lambda) \\ &= \int_0^t \int_{t-g_k}^{\infty} \lambda e^{-(y_k\lambda)} \frac{\lambda^k}{\Gamma(k)} g_k^{k-1} e^{-(g_k\lambda)} dy_k dg_k \quad \text{(draw pictures to get integration limits)} \\ &= \int_0^t e^{-(t-g_k)\lambda} \frac{\lambda^k}{\Gamma(k)} g_k^{k-1} e^{-(g_k\lambda)} dg_k \\ &= e^{-(t\lambda)} \frac{\lambda^k}{\Gamma(k)} \int_0^t g_k^{k-1} dg_k \\ &= \frac{e^{-t\lambda}(t\lambda)^k}{k!}, \end{aligned}$$

so that $X \sim \text{Poisson}(t\lambda)$. Note that $P(X = 0) = e^{t\lambda} = e^{t\lambda}(t\lambda)^0/0!$.

2. Let (X, Y) be a pair of random variables with joint pdf given by

$$f(x, y) = \frac{x}{\theta} e^{-x/\theta} \mathbf{1}(0 < y < 1/x, x > 0)$$

for some $\theta > 0$.

(a) Find $P(1 \leq X \leq 2, Y \leq 1)$.

We have

$$\begin{aligned} P(1 \leq X \leq 2, Y \leq 1) &= \int_1^2 \int_0^{1/x} \frac{x}{\theta} e^{-x/\theta} dy dx \\ &= \int_1^2 \frac{1}{\theta} e^{-x/\theta} dx \\ &= e^{-1/\theta} - e^{-2/\theta}. \end{aligned}$$

(b) Find the marginal pdf f_X of X.

We have

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \frac{x}{\theta} e^{-x/\theta} \mathbf{1}(0 < y < 1/x, x > 0) dy \\ &= \int_0^{1/x} \frac{x}{\theta} e^{-x/\theta} dy \mathbf{1}(x > 0) \\ &= \frac{1}{\theta} e^{-x/\theta} \mathbf{1}(x > 0), \end{aligned}$$

so that $X \sim \text{Exponential}(\theta)$.

(c) Find $\mathbb{E}X$.

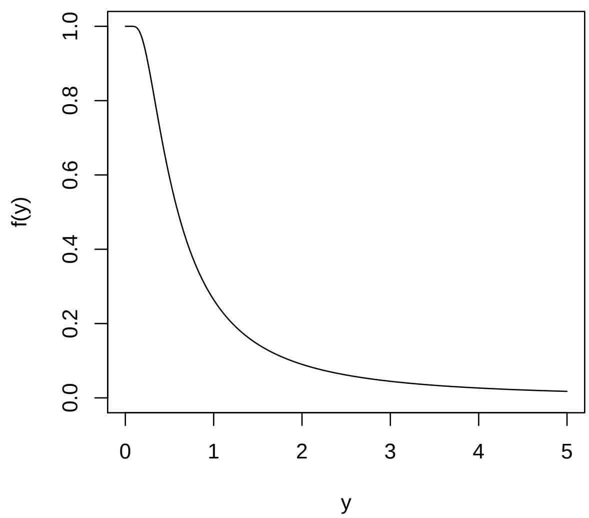
Since $X \sim \text{Exponential}(\theta)$, we have $\mathbb{E}X = \theta$.

(d) Find the marginal pdf f_Y of Y and draw a picture of it when $\theta = 1$ (you may use software).
Hint: You will have to do integration by parts.

We have

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} \frac{x}{\theta} e^{-x/\theta} \mathbf{1}(0 < y < 1/x, x > 0) dx \\ &= \int_0^{1/y} x \cdot \frac{1}{\theta} e^{-x/\theta} dx \cdot \mathbf{1}(y > 0) \\ &= \left[-x e^{-x/\theta} \Big|_0^{1/y} - \int_0^{1/y} -e^{-x/\theta} dx \right] \cdot \mathbf{1}(y > 0) \qquad \text{(by parts: } u = x,\ dv = \theta^{-1} e^{-x/\theta}) \\ &= \left[\theta - e^{-1/(y\theta)} \left(\theta + \frac{1}{y} \right) \right] \cdot \mathbf{1}(y > 0). \end{aligned}$$

With $\theta = 1$, the function looks like this:



(e) Give the conditional pdf $f(x|y)$ of $X|Y = y$ for $y = 1$ when $\theta = 1$.

For $0 < y < 1/x$ and $x > 0$ (which is the same as $0 < x < 1/y$, $y > 0$), we have

$$f(x|y) = \frac{\frac{x}{\theta}e^{-x/\theta}}{\theta - e^{-1/(y\theta)}\left(\theta + \frac{1}{y}\right)},$$

so, for $y = 1$ and $\theta = 1$, we have

$$f(x|1) = \frac{xe^{-x}}{1 - 2e^{-1}}\mathbf{1}(0 < x < 1).$$

(f) Give the conditional pdf $f(y|x)$ of $Y|X = x$ for $x > 0$.

We have

$$f(y|x) = \frac{\frac{x}{\theta}e^{-x/\theta}}{\frac{1}{\theta}e^{-x/\theta}}\mathbf{1}(0 < y < 1/x) = x \cdot \mathbf{1}(0 < y < 1/x),$$

so that $Y|X = x \sim \text{Uniform}(0, 1/x)$.

3. Let (Z_1, Z_2) be a pair of rvs with the standard bivariate Normal distribution with correlation ρ, so that their joint pdf is given by

$$f(z_1, z_2) = \frac{1}{2\pi}\frac{1}{\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right] \quad \text{for all } z_1, z_2 \in \mathbb{R}.$$

(a) Show that Z_1 and Z_2 are independent if $\rho = 0$.

If $\rho = 0$ then the joint pdf of (Z_1, Z_2) becomes

$$f(z_1, z_2) = \frac{1}{2\pi} \exp\left[-\frac{1}{2}(z_1^2 + z_2^2)\right]$$
$$= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-z_2^2/2},$$

so that it can be factored into the product of a function of only z_1 and a function of only z_2, implying independence of Z_1 and Z_2.

(b) Show that the marginal pdf of Z_1 is the $\text{Normal}(0, 1)$ distribution.

The marginal pdf f_{Z_1} is given by

$$f_{Z_1}(z_1) = \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\right] dz_2$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}\left((z_2^2 - 2\rho z_1)^2 + z_2^2 - \rho^2 z_2^2\right)\right] dz_2$$
$$= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}(z_2^2 - 2\rho z_1)^2\right] dz_2}_{=1,\text{ integral over Normal}(\rho z_1, 1-\rho^2)\text{ pdf}}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2},$$

which is the pdf of the $\text{Normal}(0, 1)$ distribution.

(c) Show that $Z_2 | Z_1 = z_1 \sim \text{Normal}(\rho z_1, 1 - \rho^2)$.

In our work towards finding the marginal pdf f_{Z_1} of Z_1, we rewrote the joint pdf of Z_1 and Z_2 as

$$f(z_1, z_2; \rho) = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-z_1^2/2}}_{f_{Z_1}(z_1)} \cdot \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}(z_2^2 - 2\rho z_1)^2\right].$$

We see from here that the conditional pdf $f(z_2|z_1)$ of $Z_2|Z_1 = z_1$ is given by

$$f(z_2|z_1) = \frac{f(z_1, z_2)}{f_{Z_1}(z_1)} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}(z_2^2 - 2\rho z_1)^2\right],$$

which is the pdf of the $\text{Normal}(\rho z_1, 1 - \rho^2)$ distribution.

4. Let X have pdf f_X and for some $\tau \in (0,1)$ define the *quantile check function* as

$$\rho_\tau(z) = z(\tau - \mathbf{1}(z<0)) = \begin{cases} z\tau, & z \geq 0 \\ -z(1-\tau), & z < 0. \end{cases}$$

(a) Show that the τ-quantile q_τ of X is equal to the value of a which minimizes $\mathbb{E}\rho_\tau(X-a)$. *Hint: Set up the integral and differentiate it with respect to a using the rule of Leibniz*

$$\frac{d}{dx}\int_{a(x)}^{b(x)} g(x,t)dt = g(x,b(x))\frac{d}{dx}b(x) - g(x,a(x))\frac{d}{dx}a(x) + \int_{a(x)}^{b(x)} \frac{d}{dx}g(x,t)dt.$$

Then show that the derivative is equal to zero when $a = q_\tau$.

We have

$$\begin{aligned}\mathbb{E}\rho_\tau(X-a) &= \int_{-\infty}^{\infty} \rho_\tau(x-a)f_X(x)dx \\ &= -(1-\tau)\int_{-\infty}^{a}(x-a)f_X(x)dx + \tau\int_a^\infty (x-a)f_X(x)dx.\end{aligned}$$

The rule of Leibniz applied to the two integrals gives

$$\begin{aligned}\frac{d}{da}\int_{-\infty}^{a}(x-a)f_X(x)dx &= -\int_{-\infty}^{a} f_X(x)dx \\ \frac{d}{da}\int_{a}^{\infty}(x-a)f_X(x)dx &= -\int_{a}^{\infty} f_X(x)dx.\end{aligned}$$

Now we write

$$\frac{d}{da}\mathbb{E}\rho_\tau(X-a) = (1-\tau)\int_{-\infty}^{a} f_X(x)dx - \tau\int_a^\infty f_X(x)dx = 0.$$

We see that $a = q_\tau$ solves the above equation since $\int_{-\infty}^{q_\tau} f_X(x)dx = \tau$ and $\int_{q_\tau}^{\infty} f_X(x)dx = 1-\tau$.

(b) Argue that the median of X is the value of a which minimizes $\mathbb{E}|X-a|$.

The median is the $\tau = 0.5$ quantile, or $q_{0.5}$. Noting that $\rho_{0.5}(X-a) = \frac{1}{2}|X-a|$, we have

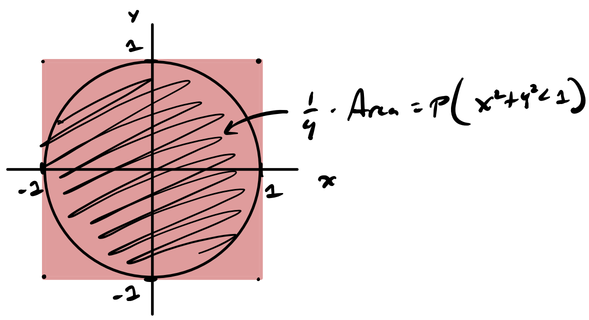
$$q_{0.5} = \underset{a}{\text{argmin}}\ \mathbb{E}\rho_{0.5}(X-a) = \underset{a}{\text{argmin}}\ \frac{1}{2}\mathbb{E}|X-a| = \underset{a}{\text{argmin}}\ \mathbb{E}|X-a|$$

5. (Optional) Additional problems from CB: 4.4, 4.5, 4.17, 4.18

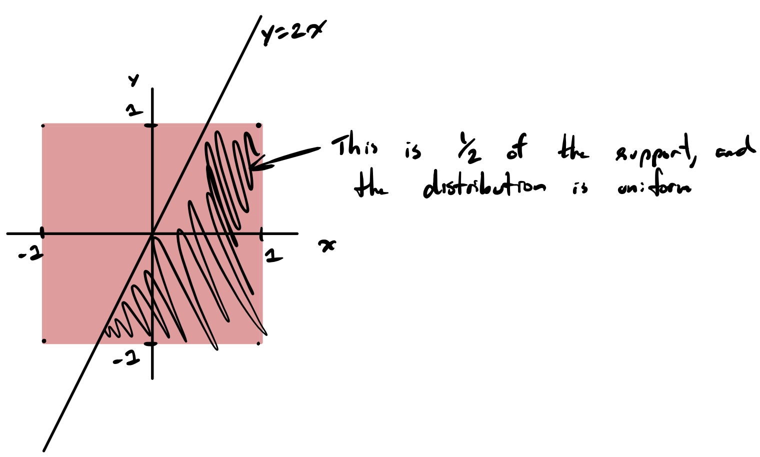
Problems 4.1, 4.9, and 4.15 from CB.

4.1 Let $(X, Y) \sim f(x,y) = \frac{1}{4} \mathbb{1}(-1 < x < 1, -1 < y < 1)$

(a) $P(X^2+Y^2 < 1) = \frac{1}{4}\pi$



(b) $P(2X - Y > 0) = P(Y < 2X) = \frac{1}{2}$

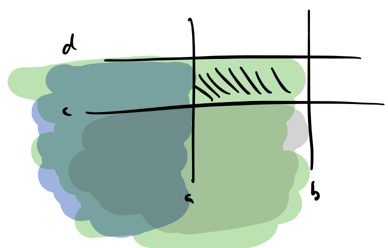


(c) $P(|X+Y| < 2) = 1$, since $|x+y| < 2 \ \forall (x,y)$ in the support.

4.9 Show

$$F(x,y) = F_X(x) F_Y(y) \implies P(a \le X \le b,\ c \le Y \le d) = P(a \le X \le b)\, P(c \le Y \le d).$$

Reasoning from the picture



we have, for any $a \le b$, $c \le d$,

$$P(a \le X \le b,\ c \le Y \le d) = P(X \le b \cap Y \le d) - P(X \le a \cap Y \le d) - P(X \le b \cap Y \le c) + P(X \le a \cap Y \le c)$$

$$= F_X(b) F_Y(d) - F_X(a) F_X(d) - F_X(b) F_Y(c) + F_X(a) F_Y(c)$$

$$= [F_X(b) - F_X(a)] F_Y(d) - [F_X(b) - F_X(a)] F_Y(c)$$

$$= [F_X(b) - F_X(a)][F_Y(d) - F_Y(c)]$$

$$= P(a \le X \le b)\, P(c \le Y \le d).$$

4.15 Let $X \sim \text{Poisson}(\theta)$, $Y \sim \text{Poisson}(\lambda)$ be indep.

We know $X+Y \sim \text{Poisson}(\lambda+\theta)$ [Can show easily with mgfs]

(i) Find dist of $X \mid X+Y$.

$$P(X=x \mid X+Y=w) = \frac{P(X=x \cap X+Y=w)}{P(X+Y=w)}$$

$$= \frac{P(X=x \cap Y=w-x)}{P(X+Y=w)}$$

$$= \frac{\dfrac{e^{-\theta}\theta^{x}}{x!}\,\dfrac{e^{-\lambda}\lambda^{w-x}}{(w-x)!}}{\dfrac{e^{-(\theta+\lambda)}(\theta+\lambda)^{w}}{w!}}$$

$$= \binom{w}{x}\left(\frac{\theta}{\theta+\lambda}\right)^{x}\left(1-\frac{\theta}{\theta+\lambda}\right)^{w-x},$$

so that $X \mid X+Y=w \sim \text{Binomial}\left(w, \frac{\theta}{\theta+\lambda}\right)$.

(ii) Following the same steps, we obtain

$$Y \mid X+Y=w \sim \text{Binomial}\left(w, \frac{\lambda}{\theta+\lambda}\right).$$