ASSIGNMENT 5

MATH 356

QUESTION 1

Qualitative

Let E_1 and E_2 be $\text{Ber}(p_1)$ and $\text{Ber}(p_2)$, respectively. $X_1 + X_2$ tallies E_1 and E_2 in a series of n trials, i.e. $X_1 + X_2$ counts $E_1 + E_2$, where $E_1 = 1$ and $E_2 = 1$ with probabilities p_1, p_2, and neither can occur ($=1$) simultaneously.

Note that $p_1 + p_2 < 1$, since the variables $E_3 \sim \text{Ber}(p_3)$ and $E_4 \sim \text{Ber}(p_4)$ also exist behind the scenes.

$\mathbb{P}(E_1 + E_2 = 1) = \mathbb{P}(E_1 = 1 \cup E_2 = 1)$, and by additivity this is $p_1 + p_2$. Thus, $X_1 + X_2$ is $\text{Bin}(n, p_1 + p_2)$.

Quantitative

$\mathbb{P}(X_1 + X_2 = k) = \sum_{i=0}^{k} \mathbb{P}(X_1 = i, X_2 = k - i)$. Using marginal probability, we have:

Note that, since $X_1 + X_2 = k$, $X_3 + X_4 = n - k$

$$\sum_{\substack{x_3, x_4 \\ x_3 + x_4 = n-k}} \sum_{i=1}^{k} \rho(i, k-1, x_3, x_4) = \sum_{\substack{x_3, x_4 \\ x_3 + x_4 = n-k}} \sum_{i=1}^{k} \binom{n}{i, k-1, x_3, x_4} p_1^i p_2^{k-1} p_3^{x_3} p_4^{x_4}$$

$$= \sum_{\substack{x_3, x_4 \\ x_3 + x_4 = n-k}} \frac{1}{k!} \binom{n}{x_3, x_4} p_3^{x_3} p_4^{x_4} \sum_{i=1}^{k} \binom{k}{i, k-1} p_1^i p_2^{k-1}$$

$$\text{By multinomial theorem} = \sum_{\substack{x_3, x_4 \\ x_3 + x_4 = n-k}} \frac{1}{k!} \binom{n}{x_3, x_4} p_3^{x_3} p_4^{x_4} (p_1 + p_2)^k$$

$$= \frac{n!}{k!(n-k)!} (p_1 + p_2)^k \sum_{\substack{x_3, x_4 \\ x_3 + x_4 = n-k}} \binom{n-k}{x_3, x_4} p_3^{x_3} p_4^{x_4}$$

$$\text{By multinomial theorem} = \frac{n!}{k!(n-k)!} (p_1 + p_2)^k (p_3 + p_4)^{n-k} = \binom{n}{k} (p_1 + p_2)^k [1 - (p_1 + p_2)]^{n-k}$$

$$\implies X_1 + X_2 \sim \text{Bin}(n, p_1 + p_2)$$

$\square$

QUESTION 2

The dy integral is $\sqrt{2\pi}$, since we are only "shifting" the normal curve. In general, $\int e^{-(x\pm a)^2} = \int e^{-x^2}$ over $\mathbb{R}$. This was discussed in class.

Part (a): One requires that $\int\int_{\mathbb{R}} f(x,y) = \int\int_{\mathbb{R}} ce^{-\frac{x^2}{2}-\frac{(x-y)^2}{2}} = 1$. Consider

$$\int\int_{\mathbb{R}} e^{-\frac{x^2}{2}-\frac{(x-y)^2}{2}} = \underbrace{\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,dx}_{=\sqrt{2\pi}}\underbrace{\int_{\mathbb{R}} e^{-\frac{(y-x)^2}{2}}\,dy}_{=\sqrt{2\pi}} = 2\pi$$

Thus, $c = \frac{1}{2\pi}$

Part (b): All we've got to do is integrate:

$$f_X(x) = \int_{\mathbb{R}} \frac{1}{2\pi}e^{-\frac{x^2}{2}-\frac{(x-y)^2}{2}}\,dy = \frac{1}{2\pi}e^{-\frac{x^2}{2}}\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2}} = \frac{\sqrt{2\pi}}{2\pi}e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

$$f_Y(y) = \int_{\mathbb{R}} \frac{1}{2\pi}e^{-\frac{x^2}{2}-\frac{(x-y)^2}{2}}\,dx = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-x^2-\frac{y^2}{2}+xy}\,dx = \frac{e^{-\frac{y^2}{2}}}{2\pi}\int_{\mathbb{R}} e^{-x^2+xy}\,dx$$

$$= \frac{e^{-\frac{y^2}{2}}}{2\pi}\int_{\mathbb{R}} e^{-(x-y/2)^2+y^2/4}\,dx = \frac{e^{-\frac{y^2}{2}+\frac{y^2}{4}}}{2\pi}\underbrace{\int_{\mathbb{R}} e^{-(x-y/2)^2}\,dx}_{=\sqrt{\pi}} = \frac{1}{2\sqrt{\pi}}e^{-\frac{y^2}{4}}$$

Part (c): We see that $f_X(x)f_Y(y) = \frac{\sqrt{2}}{4\pi}e^{-\frac{x^2}{2}-\frac{y^2}{4}}$, so $f_{X,Y} \neq f_X f_Y$, and we conclude that X and Y are not independent.

QUESTION 3

Part (a): Let $X \sim \text{Gamma}(r, \lambda)$ and $Y \sim \text{Gamma}(s, \lambda)$. Since X and Y are independent, we can write their joint probability density as follows:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} \frac{\lambda^s y^{s-1} e^{-\lambda y}}{\Gamma(s)} = \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} x^{r-1} y^{s-1} e^{-\lambda(x+y)}$$

Let $B : (x,y) \mapsto \frac{x}{x+y}$ and $G : (x,y) \mapsto x+y$. Since X and Y are defined for non-negative values (call this region K), we have that $(B,G) \in (0,1) \times (0,\infty) =: L$, where B and G are continuously differentiable on this region.

Define $q(b,g) = bg$. Then $bg = \frac{x}{x+y}(x+y) = x$.

Similarly, define $r(b,g) = g(1-b) = (x+y)\left(1 - \frac{x}{x+y}\right) = (x+y) - x = y$. We have

$$f_{B,G}(b,g) = f_{X,Y}(q,r)\,|\text{Jac}(b,g)| = \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)}(bg)^{r-1}[g(1-b)]^{s-1} e^{-\lambda g} |\text{Jac}(b,g)|$$

Lastly, $\text{Jac}(b,g) = \begin{vmatrix} g & b \\ -g & 1-b \end{vmatrix} = g(1-b) + gb = g$, so

$$f_{B,G}(b,g) = \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)}(bg)^{r-1}[g(1-b)]^{s-1} e^{-\lambda g} g \quad \text{for } (b,g) \in L \text{ , 0 otherwise}$$

Part (b): Rearranging the joint density from above, we get

$$f_{B,G}(b,g) = \left[\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} b^{r-1}(1-b)^{s-1}\right]\left[\frac{g^{r-1} g^{s-1} e^{-\lambda g} g \lambda^{r+s}}{\Gamma(r+s)}\right] = \underbrace{\left[\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} b^{r-1}(1-b)^{s-1}\right]}_{f_B \implies B\sim\text{Beta}(r,s)} \underbrace{\left[\frac{g^{r+s-1} e^{-\lambda g} \lambda^{r+s}}{\Gamma(r+s)}\right]}_{f_G \implies G\sim\text{Gamma}(r+s,\lambda)}$$

We arrive at the following with regards to part (a):

1. X, Y are independent with $X \sim \text{Gamma}(r, \lambda)$ and $Y \sim \text{Gamma}(s, \lambda)$

2a. $B = \frac{X}{X+Y}$ and $G = X + Y$ 2b. $X = BG$ and $Y = G(1-B)$

3. B, G are independent with $B \sim \text{Beta}(r,s)$ and $G \sim \text{Gamma}(r+s, \lambda)$

One sees that the variables described in part (b) are defined and distributed precisely as B, G, X, Y from part (a), and we conclude that the distribution of $(BG, G(1-B))$ is just

$$f_{X,Y}(x,y) = \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} x^{r-1} y^{s-1} e^{-\lambda(x+y)} \quad \text{for } (x,y) \in K \text{ , 0 otherwise}$$

Question 4

Let $X_1, ..., X_n$ be independent, exponentially distributed variables with parameter λ_i for X_i. By independence, their joint density is

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1)...f_{X_n}(x_n) = \lambda_1...\lambda_n e^{-x_1\lambda_1-...-x_n\lambda_n}$$

Let $Y = X_i$ be $\min(X_1, ..., X_n)$ and $I = i : X_i = Y$, defined for $1 \le i \le n$. We have

$$\begin{aligned}\mathbb{P}(I = i) &= \mathbb{P}(X_i = \min(X_1, ..., X_n)) = \mathbb{P}(X_i \le X_1, ..., X_i \le X_n) \\ &= \underset{X_i \le X_j : i \ne j}{\int ... \int} f = \int_0^\infty \int_{x_i}^\infty ... \int_{x_i}^\infty \lambda_1...\lambda_n e^{-x_1\lambda_1-...-x_n\lambda_n} dx_1...dx_n dx_i \\ &= \lambda_1...\lambda_n \int_0^\infty e^{-x_i\lambda_i} \underbrace{\int_{x_i}^\infty e^{-x_n\lambda_n} ... \int_{x_i}^\infty e^{-x_1\lambda_1}}_{n-1 \text{ times}} dx_1...dx_n dx_i \qquad \star\end{aligned}$$

For any $j \ne i$, we have

$$\int_{x_i}^\infty e^{-x_j\lambda_j} = -\frac{1}{\lambda_j}\left[e^{-x_j\lambda_j}\right]_{x_i}^\infty = \frac{e^{-x_i\lambda_j}}{\lambda_j}$$

Evaluating for the j^{th} integrand and moving leftward, as in $\star$, we see that λ_j cancels and the remaining $e^{-x_i\lambda_j}$ term may be moved into the first (dx_i) integral. We are then left with

$$\star = \lambda_i \int_0^\infty e^{-x_i(\lambda_1+...+\lambda_n)} = -\frac{\lambda_i}{\lambda_1 + ... + \lambda_n}\left[e^{-x_i(\lambda_1+...+\lambda_n)}\right]_0^\infty = \frac{\lambda_i}{\lambda_1 + ... + \lambda_n}$$

$\square$

QUESTION 5

Let X_1, X_2, X_3 be independent variables distributed as $\text{Exp}(\lambda)$. Their joint PDF is

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = \lambda^3 e^{-\lambda(x_1+x_2+x_3)} = \lambda^3 e^{-\lambda x_1} e^{-\lambda x_2} e^{-\lambda x_3}$$

Notice that f is symmetric, and so X_1, X_2, X_3 are exchangeable, and thus

$$\mathbb{P}(X_1 < X_2 < X_3) = \mathbb{P}(X_{\sigma(1)} < X_{\sigma(2)} < X_{\sigma(3)})$$

Since X_i are continuous variables, we can remove any cases of equality and still maintain an intact & equivalent probability. Also, notice that there are 6 permutations $\sigma(i)$, so in particular

$$6\mathbb{P}(X_1 < X_2 < X_3) = 1 \implies \mathbb{P}(X_1 < X_2 < X_3) = \frac{1}{6}$$

One can check manually by integrating over the constaints $X_1 < X_2$ and $X_2 < X_3$:

$$\begin{aligned}\lambda^3 \int_0^\infty e^{-\lambda x_2} dx_2 \int_{x_2}^\infty e^{-\lambda x_3} dx_3 \int_0^{x_2} e^{-\lambda x_1} dx_1 &= \lambda^2 \int_0^\infty e^{-\lambda x_2}(1 - e^{-\lambda x_2}) dx_2 \int_{x_2}^\infty e^{-\lambda x_3} dx_3 \\ &= \lambda \int_0^\infty e^{-2\lambda x_2} - e^{-3\lambda x_2} dx_2 = \frac{1}{6}\end{aligned}$$

□