Stat 512
HW 2 Solutions
October 24, 2003

- 3.18

  Part 1

  Given $X_1,\ X_2$ are iid random variables, show that you cannot satisfy $P(X_1+X_2=k)=\frac{1}{11}$ for $k=2,...,12$.

  Since X_1+X_2 is always an integer between 2 and 12, X must only take values between 1 and 6. Assume $P(X_1+X_2=k)=\frac{1}{11}$.

$$\begin{aligned} P(X_1+X_2=2) &= P(X_1=1 \ \& \ X_2=1) \\ &= P(X_1=1)\cdot P(X_2=1) \\ &= P(X=1)^2=\frac{1}{11} \\ P(X=1) &= \sqrt{\frac{1}{11}} \end{aligned}$$

  Similarly,

$$\begin{aligned} P(X_1+X_2=12) &= P(X_1=6 \ \& \ X_2=6) \\ &= P(X_1=6)\cdot P(X_2=6) \\ &= P(X=6)^2=\frac{1}{11} \\ P(X=6) &= \sqrt{\frac{1}{11}} \end{aligned}$$

  Now for when $k=7$,

$$\begin{aligned} P(X_1+X_2=7) &\geq P(X_1=1 \ \& \ X_2=6)+P(X_1=6 \ \& \ X_2=1) \\ &\geq P(X=1)\cdot P(X=6)+P(X=6)\cdot P(X=1) \\ &\geq \frac{1}{11}+\frac{1}{11}=\frac{2}{11} \\ \frac{1}{11} &< \frac{2}{11} \end{aligned}$$

  So $P(X_1+X_2=7)\neq\frac{1}{11}$. This corresponds exactly to the case where we have two identical dice and want the sum of their faces to be uniformly distributed.

  Part 2

  Is it possible to weight a pair of dice to get every sum from 2 to 12 with the same probability?

No, follow the same steps as Part 1, but use p_1 and q_1, with p_6 and q_6.

Now $p_1 \cdot q_1 = \frac{1}{11}$ & $p_6 \cdot q_6 = \frac{1}{11}$
then $p_1 \cdot q_6 + p_6 \cdot q_1 > \frac{1}{11}$ since either $p_1 \cdot q_6 > \frac{1}{11}$, $p_6 \cdot q_1 > \frac{1}{11}$, or they're both $\frac{1}{11}$.

- 3.4.8
  Given $X \sim B(m,p)$, $Y \sim B(n,p)$ and letting $Z = X + Y$, find $P(X = x|Z = z)$.

$$
\begin{aligned}
P(X = x|Z = z) &= \frac{P(X = x \;\&\; Z = z)}{P(Z = z)} \\
&= \frac{P(X = x \;\&\; Y = z - x)}{P(Z = z)} \\
&= \frac{P(X = x) \cdot P(Y = z - x)}{P(Z = z)} \\
&= \frac{\binom{m}{x} p^x (1-p)^{(m-x)} \binom{n}{z-x} p^{(z-x)} (1-p)^{(n-(z-x))}}{\sum_{i=0}^{z} \binom{m}{i} p^i (1-p)^{(m-i)} \binom{n}{z-i} p^{(z-i)} (1-p)^{(n-(z-i))}} \\
&= \frac{\binom{m}{x}\binom{n}{z-x} p^z (1-p)^{(n+m-z)}}{\sum_{i=0}^{z} \binom{m}{i}\binom{n}{z-i} p^z (1-p)^{(m+n-z)}} \\
&= \frac{\binom{m}{x}\binom{n}{z-x}}{\sum_{i=0}^{z} \binom{m}{i}\binom{n}{z-i}} \\
&= \frac{\binom{m}{x}\binom{n}{z-x}}{\binom{m+n}{z}}
\end{aligned}
$$

So $P(X = x|Z = z)$ is Hypergeometric.

- 3.5.2
  Given $(X_1, X_2, X_3, X_4, X_5) \sim$ Multinomial $(p_1, p_2, p_3, p_4, p_5, n = 30)$, what is the distribution of $(X_1, X_2, X_3, X_4|X_5 = 2)$?

$$
P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4|X_5 = 2) = \frac{P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = 2)}{P(X_5 = 2)}
$$

$$
\begin{aligned}
&= \frac{\binom{30}{x_1, x_2, x_3, x_4, 2} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4} p_5^2}{\binom{30}{2} p_5^2 (1-p_5)^{28}} \\
&= \frac{\frac{30!}{x_1! x_2! x_3! x_4! 2!} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4}}{\frac{30!}{2!28!} (1-p_5)^{(x_1+x_2+x_3+x_4)}} \\
&= \frac{28!}{x_1! x_2! x_3! x_4!} \left(\frac{p_1}{1-p_5}\right)^{x_1} \left(\frac{p_2}{1-p_5}\right)^{x_2} \left(\frac{p_3}{1-p_5}\right)^{x_3} \left(\frac{p_4}{1-p_5}\right)^{x_4}
\end{aligned}
$$

This is exactly a Multinomial($\frac{p_1}{1-p_5}, \frac{p_2}{1-p_5}, \frac{p_3}{1-p_5}, \frac{p_4}{1-p_5}, n = 28$).

- 3.26
  Assuming $P(X = x) = (1-p)p^x$:

$$
\begin{aligned}
P(X \geq x) &= \sum_{i=x}^{\infty} (1-p)p^i \\
&= p^x (1-p) \sum_{i=0}^{\infty} p^i \\
&= p^x (1-p) \frac{1}{1-p} = p^x
\end{aligned}
$$

$$
\begin{aligned}
P(X \geq j+k | X \geq k) &= \frac{P(X \geq j+k \ \& \ X \geq k)}{P(X \geq k)} \\
&= \frac{P(X \geq j+k)}{P(X \geq k)} \\
&= \frac{p^{j+k}}{p^k} = p^j = P(X \geq j)
\end{aligned}
$$

- 4.2.9
  Given: (X_1, X_2, X_3) are independent uniform r.v.'s with $f_{X_i}(x) = \frac{1}{(b-a)}$ for $a < x < b$ and $(i = 1, 2, 3)$.

  (a) Find the median of X_i.
  Since each X_i is identically distributed, the median will be the same for $i = 1, 2, 3$ and since each X_i is $\mathcal{U}(a, b)$, then the center will be the median, $\frac{(a+b)}{2}$.

  (b) Let $Y = min(X_1, X_2, X_3)$, and $a < c_i < b$ for $i = 1, 2$. Compute $P(Y > c_1)$ and $P(Y < c_2)$.

$$\begin{aligned}
P(min > c_1) &= P(\text{all } X_i > c_1) \\
&= P(X > c_1)^3 \\
&= \left(\frac{b - c_1}{b - a}\right)^3 \\
P(min < c_2) &= 1 - P(min > c_2) \\
&= 1 - \left(\frac{b - c_2}{b - a}\right)^3
\end{aligned}$$

(c) Let $U = max(X_1, X_2, X_3)$ and $a < c_i < b$ for $i = 1, 2$. Compute $P(U > c_1)$ and $P(U < c_2)$. Similarly,

$$\begin{aligned}
P(max < c_2) &= P(\text{all } X_i < c_2) \\
&= P(X < c_2)^3 \\
&= \left(\frac{c_2 - a}{b - a}\right)^3 \\
P(max > c_1) &= 1 - P(max < c_1) \\
&= 1 - \left(\frac{c_1 - a}{b - a}\right)^3
\end{aligned}$$

(d) If $c_1 < c_2$, then compute $P(U < c_2, Y > c_1)$.

$$\begin{aligned}
P(c_1 < min \ \& \ max < c_2) &= P(X_1, X_2, X_3 \in (c_1, c_2)) \\
&= P(X_i \in (c_1, c_2))^3 \\
&= \left(\frac{c_2 - c_1}{b - a}\right)^3
\end{aligned}$$

- 4.7
  Let $f_{X,Y}(x, y) = e^{-(x+y)}$, $x, y > 0$.

  (a) Find $P(X > 1)$.

$$\begin{aligned}
P(X > 1) &= \int_1^\infty \int_0^\infty e^{-(x+y)} \, dy \, dx \\
&= e^{-1} = \frac{1}{e}
\end{aligned}$$

  (b) Find $P(a < X + Y < b)$.
  There are two methods of doing this:

Method 1: Use only the bounds of integration:

$$\begin{aligned} P(a < X+Y < b) &= \int_0^b \int_{(a-x)^+}^{b-x} e^{-(x+y)}\,dy\,dx \\ &= \int_0^a \int_{a-x}^{b-x} e^{-(x+y)}\,dy\,dx + \int_a^b \int_0^{b-x} e^{-(x+y)}\,dy\,dx \\ &= a(e^{-a} - e^{-b}) - e^{-b}(b-a) + (e^{-a} - e^{-b}) \\ &= \frac{a+1}{e^a} - \frac{b+1}{e^b} \end{aligned}$$

Method 2: Recognize that X and Y are independent exponential(1) random variables, so $X + Y = Z \sim Gamma(1,1)$ (by Scott's notation).

$$\begin{aligned} P(a < Z < b) &= \int_a^b z e^{-z}\,dz \\ &= \frac{a+1}{e^a} - \frac{b+1}{e^b} \end{aligned}$$

(c) Find $P(X < Y | X < 2Y)$:

$$\begin{aligned} P(X < Y|X < 2Y) &= \frac{P(X < Y \ \&\ X < 2Y)}{P(X < 2Y)} \\ &= \frac{\int_0^\infty \int_0^y e^{-(x+y)}\,dx\,dy}{\int_0^\infty \int_0^{2y} e^{-(x+y)}\,dx\,dy} \\ &= \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4} \end{aligned}$$

- 4.14 Let $\mathbf{X} = (X_1, X_2)$ have the joint density function

$$f_{X_1,X_2}(x_1,x_2) = f_1(x_1)f_2(x_2)[1 + \alpha[2F_1(x_1) - 1][2F_2(x_2) - 1]],\ |\alpha| \leq 1 \text{ where}$$

$f_i(x_i)$ and $F_i(x_i)$ are the pdf and cdf of a $\mathcal{N}(0,1)$

(a) Verify that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2 = 1$
First, a few definitions to simplify the algebra. Let $H(x) = 2F(x) - 1$, where $F(x)$ is your favorite of the two F_is, since they are identical. Similarly, let f be your favorite of the two f_is.

Step 1: Show H(x) is odd and f(x) is even.

$$f(-x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(-x)^2}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\
&= f(x) \\
H(-x) &= 2F(-x) - 1 \\
F(-x) &= \int_{-\infty}^{-x} f(t)\, dt \\
&= 1 - \int_{-x}^{\infty} f(t)\, dt \\
&= 1 - \int_{-\infty}^{x} f(t)\, dt \quad \text{(since f(t) is even, and a pdf)} \\
&= 1 - F(x) \\
H(-x) &= 2(1 - F(X)) - 1 \\
&= 1 - F(x) \\
&= -H(x)
\end{aligned}
$$

Step 2: Show an odd function integrated over $\Re$ is zero:
Let $g(-x) = -g(x)$,

$$
\begin{aligned}
\int_{-\infty}^{\infty} g(x)\, dx &= \int_{-\infty}^{0} g(x)\, dx + \int_{0}^{\infty} g(x)\, dx \\
&= -\int_{0}^{\infty} g(x)\, dx + \int_{0}^{\infty} g(x)\, dx \\
&= 0
\end{aligned}
$$

Step 3: Do the integration, recalling that an even function times an odd function is odd.

$$
\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_1\, dx_2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1)f(x_2)[1 + \alpha H(x_1)H(x_2)]\, dx_1\, dx_2 \\
&= \int_{-\infty}^{\infty} f(x_2)\left[\int_{-\infty}^{\infty} f(x_1)\, dx_1 + \alpha H(x_2)\int_{-\infty}^{\infty} f(x_1)H(x_1)\, dx_1\right] dx \\
&= \int_{-\infty}^{\infty} f(x_2)[1 + \alpha H(x_2)\cdot 0]\, dx_2 \\
&= 1
\end{aligned}
$$

Alternate method: Use integration by parts to reduce things.

(b) Find $f_{X_i}(x_i)$:
By symmetry, $f_{X_1}(x_1) = f_{X_2}(x_2)$. Now integrating as in the previous part:

$$
f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)\, dx_1
$$

$$
\begin{aligned}
&= f(x_2)\left[\int_{-\infty}^{\infty} f(x_1)\,dx_1 + \alpha H(x_2)\int_{-\infty}^{\infty} f(x_1)H(x_1)\right]dx_1 \\
&= f(x_2)\left[1 + \alpha H(x_2)\cdot 0\right] \\
&= f(x_2) = f_2(x_2)
\end{aligned}
$$

So $X_i \sim \mathcal{N}(0,1)$ marginally.

(c) Is (X_1,X_2) jointly normal?
If it is, then it will follow the form of the bivariate normal distribution:

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\; e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

This can be handled in two cases: $\alpha = 0$ and $\alpha \neq 0$.

When $\alpha = 0$ then the joint is defined as being factored into the product of two univariate pdf's. Since these pdf's are normal as well, the joint distribution must be normal because it's the product of two normals. This is also easily seen if the univariate pdf is multiplied out, then you'll have a form of the bivariate normal pdf as shown below (with $\mu_1 = \mu_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$, and $\rho = 0$).

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi}e^{-\frac{1}{2}\left[x_1^2+x_2^2\right]}$$

When $\alpha \neq 0$, we're still looking for $f_{X_1,X_2}(x_1,x_2)$ to take a form of the bivariate normal density for (X_1,X_2) to be jointly normal.
Thus $\mu_1 = \mu_2 = 0$, and $\sigma_1^2 = \sigma_2^2 = 1$. Now this bivariate form must hold for all (x_1,x_2), and examinging the case when $(x_1,x_2) = (0,0)$ results in $f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi}$. In the general bivariate normal case, the result is $\frac{1}{2\pi\sqrt{1-\rho^2}}$, thus ρ must $= 0$.
The last part we need to consider is the $H(x_i)$ stuff: $[1+\alpha(2F_1(x_1)-1)(2F_2(x_2)-1)] = 1$ or $[\alpha(2F_1(x_1)-1)(2F_2(x_2)-1)] = 0$ for all (x_1,x_2). This is not the case, thus (X_1,X_2) is not jointly normal when $\alpha \neq 0$.

- 4.17
Show "lack of memory" property of the exponential r.v.
$X \sim \mathcal{E}(\lambda)$ means $f_X(x) = \lambda e^{-\lambda x}$, $x > 0$

$$
\begin{aligned}
P(X \geq x_i) &= \int_{x_i}^{\infty} \lambda e^{-\lambda x}\,dx \\
&= e^{-\lambda x_i}
\end{aligned}
$$

$$
\begin{aligned}
P(X \geq x_1 + x_2 | X \geq x_1) &= \frac{P(X \geq x_1 + x_2 \ \& \ X \geq x_1)}{P(X \geq x_1)} \\
&= \frac{e^{-\lambda(x_1+x_2)}}{e^{-\lambda x_1}} \\
&= e^{-\lambda x_2} \\
&= P(X \geq x_2)
\end{aligned}
$$

Alternatively using the cdf:

$$
\begin{aligned}
P[X \geq x_1 + x_2 | X \geq x_1] &= \frac{P[X \geq x_1 + x_2 \text{ and } X \geq x_1]}{P[X \geq x_1]} \\
&= \frac{P[X \geq x_1 + x_2]}{P[X \geq x_1]} \\
&= \frac{(1 - P[X \leq x_1 + x_2])}{(1 - P[X \leq x_1])} \\
&= \frac{(1 - (1 - e^{-\lambda(x_1+x_2)}))}{(1 - (1 - e^{-\lambda x_1}))} \qquad \text{(exponential cdf)} \\
&= \frac{e^{-\lambda(x_1+x_2)}}{e^{-\lambda x_1}} = e^{-\lambda x_2} = P[X \geq x_2]
\end{aligned}
$$

- 4.22

  If X and Y have joint density function:

$$
f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{1}{4\pi}x^2 - y}, \ -\infty < x < \infty, \ 0 < y < \infty
$$

  Find $f_X(x)$ and $f_Y(y)$, and determine if X and Y are independent.

$$
\begin{aligned}
f_X(x) &= \int_0^\infty \frac{1}{2\pi} e^{-\frac{1}{4\pi}x^2 - y} \, dy \\
&= \frac{1}{2\pi} e^{-\frac{x^2}{4\pi}} \int_0^\infty e^{-y} \, dy \\
&= \frac{1}{2\pi} e^{-\frac{x^2}{4\pi}} \\
f_Y(y) &= \int_{-\infty}^\infty \frac{1}{2\pi} e^{-\frac{1}{4\pi}x^2 - y} \, dx \\
&= e^{-y} \int_{-\infty}^\infty \frac{1}{2\pi} e^{-\frac{x^2}{4\pi}} \, dx \\
&= e^{-y}
\end{aligned}
$$

  Recognize the distribution of $f_X(x)$ as $\mathcal{N}(0, 2\pi)$ and the ditribution of $f_Y(y)$ as $\mathcal{E}(1)$. Technically, we did not need to do the second integration since the joint distribution factored into two distributions, but it is done here for clarity, and to remind about the definition of marginal distribution. Since the joint distribution is the product of the marginals, X and Y are independent.