Answers to the 1st problem set

1. Page 308, #3. The graphs that answer this question will be handed out in class rather than displayed here.

2. Page 308, #4.

   Answer: To show something has a beta distribution, it is enough to show that its density function is of the right form. First, we have:

$$\begin{aligned} f_{1-X}(x) &= \frac{d}{dx}F_{1-X}(x) = \frac{d}{dx}\,\mathbf{Pr}(1-X \le x) = \frac{d}{dx}\,\mathbf{Pr}(X \ge 1-x) \\ &= \frac{d}{dx}(1-\mathbf{Pr}(X \le 1-x)) = \frac{-d}{dx}\,\mathbf{Pr}(X \le 1-x) = \frac{-d}{dx}F_X(1-x) \\ &= f_X(1-x) \qquad \text{(The chain rule canceled the minus sign).} \end{aligned}$$

   In other words, take the density function of X and plug in $1-x$ in place of x, to get the density of $1-X$:

$$\begin{aligned} f_X(x) &= [\text{constant}] \cdot x^{\alpha-1}(1-x)^{\beta-1} \\ f_{1-X}(x) &= [\text{constant}] \cdot (1-x)^{\alpha-1}(1-(1-x))^{\beta-1} \\ &= [\text{constant}] \cdot x^{\beta-1}(1-x)^{\alpha-1}, \end{aligned}$$

   and that is the desired density.

3. On page 303 of DeGroot & Schervish, we read that

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

   This identity is true if α and β are **any** positive numbers.

   $\alpha+3$ is a positive number.

   Therefore this identity is true if $\alpha+3$ is put in place of α.

   Use that fact to find $\mathbf{E}\left(X^3\right)$ if $X \sim \text{Beta}(\alpha,\beta)$. By using identities satisfied by the gamma function, simplify the result so that the gamma function is not mentioned in your bottom-line answer.

   Answer: As a corollary of the identity $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ we can get:

$$\Gamma(\alpha+3) = (\alpha+2)\Gamma(\alpha+2) = (\alpha+2)(\alpha+1)\Gamma(\alpha+1) = (\alpha+2)(\alpha+1)\alpha\Gamma(\alpha),$$

   and similarly

$$\begin{aligned} \Gamma(\alpha+\beta+3) &= (\alpha+\beta+2)\Gamma(\alpha+\beta+2) = (\alpha+\beta+2)(\alpha+\beta+1)\Gamma(\alpha+\beta+1) \\ &= (\alpha+\beta+2)(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta). \end{aligned}$$

We can apply these as follows:

$$\begin{aligned}
\mathbf{E}(X^3) &= \int_{-\infty}^{\infty} x^3 f_X(x)\,dx \overset{\substack{\text{Because } f_X(x)=0 \\ \text{except when } 0<x<1.\\ \downarrow}}{=} \int_0^1 x^3 f_X(x)\,dx = \int_0^1 x^3 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+3)-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+3)\Gamma(\beta)}{\Gamma(\alpha+3+\beta)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{(\alpha+2)(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta+2)(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)} \\
&= \frac{(\alpha+2)(\alpha+1)\alpha}{(\alpha+\beta+2)(\alpha+\beta+1)(\alpha+\beta)}.
\end{aligned}$$

4. Page 320, #1.

Answer: When $\alpha = \beta = 1$ then we have

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(2)}{\Gamma(1)\Gamma(1)} x^0(1-x)^0 = 1$$

if $0 < x < 1$, and $f_X(x) = 0$ if $x < 0$ or $x > 1$. In short

$$f_X(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{if } x < 0 \text{ or } x > 1 \end{cases}$$

so that the random variable X is uniformly distributed on the interval $[0, 1]$. We have

$$\begin{aligned} P &\sim \text{Uniform}(0,1), \\ X \mid P = p &\sim \text{Bin}(n,p). \end{aligned}$$

Consequently for any $x \in \{0, 1, 2, 3, \ldots, n\}$ we have

$$\begin{aligned}
\mathbf{Pr}(X = x) &\overset{\substack{\text{Law of Total Probability}\\ \downarrow}}{=} \mathbf{E}(\mathbf{Pr}(X = x \mid P)) = \mathbf{E}\left(\binom{n}{x} P^x (1-P)^{n-x}\right) \\
&= \int_0^1 \binom{n}{x} p^x (1-p)^{n-x} f_P(p)\,dp = \int_0^1 \binom{n}{x} p^x (1-p)^{n-x}\,dp \\
&= \binom{n}{x} \int_0^1 p^x (1-p)^{n-x}\,dp = \binom{n}{x} \frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)} \\
&= \frac{n!}{x!(n-x)!} \cdot \frac{x!(n-x)!}{(n+1)!} = \frac{1}{n+1}.
\end{aligned}$$

So all $n+1$ of the outcomes have the same probability; X is uniformly distributed.

5. On page 314 of DeGroot & Schervish, the authors appear to assume the reader does not know matrix algebra. For any matrix A, let A' be its transpose, so that in particular

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}' = \begin{bmatrix} x_1, x_2 \end{bmatrix}.$$

Let $f(x_1, x_2)$ be just as on page 313 of DeGroot & Schervish. Let $V = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$;

this 2×2 matrix can be considered the variance of the random vector $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$.

(a) Show that $f(x_1, x_2) = \frac{1}{2\pi} \cdot \frac{1}{\sqrt{\det V}} \cdot \exp\left\{ \frac{-1}{2} \cdot (x-\mu)' V^{-1} (x-\mu) \right\}$ where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Answer: First, observe that the inverse of V is

$$\begin{aligned} V^{-1} &= \frac{1}{\det V} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \\ &= \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/(\sigma_1\sigma_2) \\ -\rho/(\sigma_1\sigma_2) & 1/\sigma_2^2 \end{bmatrix}. \end{aligned}$$

Therefore we have

$$\begin{aligned} & (x-\mu)' V^{-1} (x-\mu) \\ = & \frac{1}{1-\rho^2} [x_1-\mu_1,\ x_2-\mu_2] \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} \\ = & \frac{1}{1-\rho^2} [x_1-\mu_1,\ x_2-\mu_2] \begin{bmatrix} \frac{x_1-\mu_1}{\sigma_1^2} - \rho\frac{x_2-\mu_2}{\sigma_1\sigma_2} \\ \frac{x_2-\mu_2}{\sigma_2^2} - \rho\frac{x_1-\mu_1}{\sigma_1\sigma_2} \end{bmatrix} \\ = & \frac{1}{1-\rho^2} \left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_1-\mu_2)^2}{\sigma_2^2} - \rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} \right) \\ = & \frac{1}{1-\rho^2} \left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right] \end{aligned}$$

and $-1/2$ times this is just what is inside the exponential function in **(5.12.4)** on page 313 of DeGroot & Schervish. Next we deal with $\dfrac{1}{\sqrt{\det V}}$:

$$\det V = \det \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = (1-\rho^2)\sigma_1^2\sigma_2^2.$$

So $\dfrac{1}{\sqrt{\det V}} = \dfrac{1}{(1-\rho^2)^{1/2}\sigma_1\sigma_2}$, and that expression preceeds the exponential function in **(5.12.4)**.

(b) Rely on **(5.12.8)** on page 317; use the same notation and assume that what it says there is true; don't try to prove it from scratch. Let $A = [a_1, a_2]$. Let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$. Show that $Y = AX + b$. Show that $\mathbf{E}(Y) = A\,\mathbf{E}(X) + b$. Show that $\mathbf{var}(Y) = AVA'$.

<u>**Answer:**</u>

$$\begin{aligned} AVA' &= [a_1,\ a_2] \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = [a_1,\ a_2] \begin{bmatrix} a_1\sigma_1^2 + a_2\rho\sigma_1\sigma_2 \\ a_1\rho\sigma_1\sigma_2 + a_2\sigma_2^2 \end{bmatrix} \\ &= (a_1^2\sigma_1^2 + a_1a_2\rho\sigma_2\sigma_2) + (a_1a_2\rho\sigma_2\sigma_2 + a_2^2\sigma_2^2) \\ &= a_1^2\sigma_1^2 + 2a_1a_2\rho\sigma_2\sigma_2 + a_2^2\sigma_2^2 = \underset{\substack{\uparrow \\ \boxed{\text{By } \mathbf{(5.12.4)}}}}{}\mathbf{var}(Y). \end{aligned}$$

This is of course the natural generalization of an identity you have learned, about variances of <u>scalar</u>-valued random variables: If U is a univariate random variable and a is a constant, then $\mathbf{var}(aU) = a^2\,\mathbf{var}(U)$. In the bivariate case, instead of pulling out a^2, we pull out A on the left and A' on the right.

(The next two problems are less "theoretical" and more "applied" than the foregoing.)

6. Page 318, # 1.

Answer: The information given on page 317 of DeGroot & Schervish can be more tersely summarized thus:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} 66.8 \\ 70 \end{bmatrix}, \begin{bmatrix} 2^2 & (0.68)(2)(2) \\ (0.68)(2)(2) & 2^2 \end{bmatrix} \right).$$

Applying **(5.12.6)** on page 315, we get

$$\begin{aligned} \mathbf{E}(X \mid Y = 72) &= 66.8 + (0.68)(2)\left(\frac{72-70}{2}\right) = 68.16, \\ \text{and } \mathbf{var}(X \mid Y = 72) &= (1 - 0.68^2) \cdot 2^2 = 2.1504 \cong 1.46642422^2. \\ \text{So } X \mid [Y = 72] &\sim N(68.16,\ 1.46642422^2). \\ \text{This implies that} & \\ W = \left.\frac{X - 68.16}{1.46642422}\right| [Y = 72] &\sim N(0,\ 1), \\ \text{and therefore} & \\ \mathbf{Pr}(W \le \underbrace{1.645}_{\substack{\uparrow \\ \text{From the table} \\ \text{on page 778.}}} \mid Y = 72) &= 0.95. \end{aligned}$$

When $W = 1.645$ then $X = 68.16 + (1.645)(1.46642422) \cong 70.572$. So 70.572 is the quantile that we sought.

7. In **Example 5.12.4** on pages 317-318, find the probability that the sum of the heights of the husband and the wife exceeds 140 inches.

Answer: To find $\mathbf{SD}(X+Y)$ we can use **(5.12.8)**: We have X and Y instead of X_1 and X_2, and we have $a_1 = a_2 = 1$. Alternatively, we can just say

$$\mathbf{var}(X+Y) = \mathbf{var}(X) + 2\,\mathbf{cov}(X,Y) + \mathbf{var}(Y) = 2^2 + 2(0.68)(2)(2) + 2^2 = 13.44 \cong 3.66606^2.$$

To find $\mathbf{E}(X+Y)$, just add: $\mathbf{E}(X+Y) = \mathbf{E}(X) + \mathbf{E}(Y) = 66.8 + 70 = 136.8$. Then

$$\mathbf{Pr}(X+Y > 140) = \mathbf{Pr}\left(\frac{X+Y-136.8}{3.66606} > \frac{140-136.8}{3.66606}\right) = P(Z > 0.8728\ldots) \cong 0.19.$$