Practice Problem 1 There are $\binom{12}{4}$ ways to select 4 of the 12 animals for group A, $\binom{8}{4}$ ways to select 4 of the remaining 8 animals for group B, and $\binom{4}{4}$ ways to select 4 of the remaining 4 animals for group C. By the multiplication rule, we have

$$\begin{aligned}\binom{12}{4}\binom{8}{4}\binom{4}{4} &= \frac{12!}{4!4!4!} \\ &= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5}{12 \times 8 \times 6} \\ &= (11 \times 10 \times 9)(7 \times 5) \\ &= (1000 - 10)35 \\ &= 34650.\end{aligned}$$

We may also recognize this as a multinomial and evaluate

$$\binom{12}{4,4,4} = \frac{12!}{4!4!4!} = 34650.$$

For the exam, it would be acceptable to leave this in the form of $(11 \times 10 \times 9 \times 7 \times 5)$.

Practice Problem 2 Let L_1 denote the event that litter 1 is chosen, and let L_2 denote the event that litter two is chosen. Both events have probability 1/2, that is

$$P(L_1) = P(L_2) = 1/2.$$

Let B denote the event that the chosen mouse has brown hair. The probability that a mouse chosen from litter 1 has brown hair is

$$P(B|L_1) = 2/3,$$

and the probability that a mouse chosen from litter 2 has brown hair is

$$P(B|L_2) = 3/5.$$

The probability that the chosen mouse has brown hair is

$$P(B) = P(B|L_1)P(L_1) + P(B|L_2)P(L_2) = (2/3 + 3/5)/2 = 19/30.$$

Given that a brown-haired mouse was selected, the probability that it came from litter 1 is

$$P(L_1|B) = \frac{P(B|L_1)P(L_1)}{P(B)} = \frac{(2/3)(1/2)}{19/30} = 10/19.$$

Practice Problem 4 $X \sim \text{Poi}(\lambda)$ has distribution function

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!},\ x = 0, 1, \ldots,$$

and $Y|(X = k) \sim \text{Bin}(k, p)$ has distribution function

$$P(Y = y|X = k) = \binom{k}{y} p^k (1-p)^{n-k},\ y = 1, \ldots, k.$$

a) $P(X = 1, Y = 0) = P(Y = 0|X = 1)P(X = 1) = (1 - p)\lambda e^{-\lambda}$.

b) $P(X = 1, Y = 0) = P(Y = 2|X = 1)P(X = 1) = 0$.

c) $\mathrm{E}Y = \mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(pX) = p\mathrm{E}X = p\lambda$.

d) No, the conditional distribution of Y given X changes depending on what value X realizes.

Practice Problem 5

a) $P(X = 2) = 0$ because $F(x)$ is continuous at 2.

b) $P(X = 0) = P(X \leq 0) - P(X < 0) = 1 - 0.9e^0 - 0 = 0.1$.

c) $P(X > 2) = 1 - P(X \leq 2) = 1 - F(2) = 0.9e^{-2}$.

Practice Problem 6 The mean and variance of a $\mathcal{U}(0, 1)$ random variable are $1/2$ and $1/12$. Since X, Y, and Z have the same distribution, they all have the same mean and variance. Since they are independent, their covariances are zero.

a) $\mathrm{E}(2X + Y - Z) = 2\mathrm{E}X + \mathrm{E}Y - \mathrm{E}Z = 1$.

b) $\mathrm{Var}(2X + Y - Z) = 4\mathrm{Var}X + \mathrm{Var}Y + \mathrm{Var}Z = 1/4$.

c)

$$\begin{aligned}
\mathrm{Cov}(2X + Y - Z, X - Z) &= \mathrm{Cov}(2X + Y - Z, X) - \mathrm{Cov}(2X + Y - Z, Z) \\
&= 2\mathrm{Var}X + \mathrm{Cov}(Y, X) - \mathrm{Cov}(Z, X)] \\
&\quad - 2\mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z) + \mathrm{Var}Z \\
&= 2\mathrm{Var}X + \mathrm{Var}Z \\
&= 1/4.
\end{aligned}$$

Practice Problem 7

a) For $x > 10$, the cdf is

$$\begin{aligned}
F_X(x) &= \int_{10}^{x} 10/t^2 \, dt \\
&= 10 \int_{10}^{x} t^{-2} \, dt \\
&= 10 \left[-\frac{1}{t} \right]_{10}^{x} \\
&= 1 - \frac{10}{x}.
\end{aligned}$$

The full cdf is

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 10, \\ 1 - \frac{10}{x} & \text{if } x > 10. \end{cases}$$

Solving

$$1/2 = F_X(x) = 1 - \frac{10}{x}$$

indicates that the median is 20.

b) Since X has positive density for $0 < x < \infty$, $Y = 1/X$ has positive density for $0 < y < 1/10$. That is, the cdf of Y will be zero for $y \leq 0$ and one for $y \geq 1/10$. For $0 < y < 1/10$, $1/y > 10$, and the cdf is

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(1/X \leq y) \\
&= P(X \geq 1/y) \\
&= 1 - P(X < 1/y) \\
&= 1 - P(X \leq 1/y) \\
&= 1 - F_X(1/y) \\
&= 1 - (1 - 10y) \\
&= 10y.
\end{aligned}$$

The full cdf is

$$F_X(x) = \begin{cases} 0 & \text{if } y \le 0, \\ 10y & \text{if } 0 < y < 1/10, \\ 1 & \text{if } y \ge 1/10. \end{cases}$$

c) The probability that a device functions for more than 15 hours is

$$p = P(X < 15) = 1 - F_X(15) = 2/3.$$

For $i = 1, \ldots, 1000$, let $W_i \sim \text{Bern}(p)$ indicate that device i functions for more than 15 hours. The sum $Z = \sum_{i=1}^{1000} W_i$ denotes the number of devices that function for more than 15 hours, and it has mean

$$\mathrm{E}Z = \mathrm{E}\sum_{i=1}^{1000} W_i = \sum_{i=1}^{1000} \mathrm{E}W_i = 1000(2/3).$$

Practice Problem 8 For $X \sim N(4, 4)$, to find $P(0 < X < 8)$, we standardize by subtracting the mean (4) and dividing by the standard deviation $\sqrt{4} = 2$,

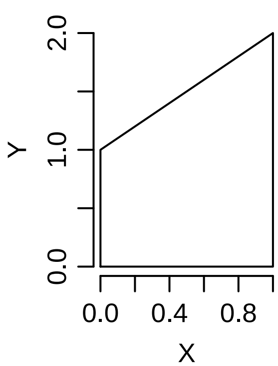
$$P(0 < X < 8) = P\left(\frac{0-4}{2} < \frac{X-4}{2} < \frac{8-4}{2}\right) = P(-2 < Z < 2) = \Phi(2) - \Phi(-2) = .9772 - .0228 = .9544$$

The distribution of $2(X-1)$ will be a normal, because a linear combination of a normal is also normal; we then just need to find the mean and variance. By the rules of expectation, $E(2X-2) = 2EX - 2 = 2(4) - 2 = 6$, and $\mathrm{Var}(2X-2) = 2^2\,\mathrm{Var}(X) = 4 * 4 = 16$, so $2(X-1) \sim N(6, 16)$.

Then since $(X-4)/2)$ is known to be standard normal, $Y = ((X-4)/2)^2$ is chi-squared with one degree of freedom. To find $P(Y < 4)$, we look up $P(Y > 4)$ in Table Vb (p662), it is .045, so $P(Y < 4) = 1 - P(Y > 4) = 1 - .045 = .955$.

(Think about this: did you expect those two probabilities to be the same?)

Practice Problem 9 X and Y are uniform on the region of $0 < X < 1$ and $0 < Y < X + 1$. A picture of this region is



So given $Y = 1/2$, X can be between 0 and 1; since it's uniform everywhere, this means $X|(Y = 1/2) \sim \mathcal{U}(0, 1)$.

But given $Y = 3/2$, can only be between 1/2 and 1, so $X|(Y = 3/2) \sim \mathcal{U}(1/2, 1)$.

To find the marginal distribution for Y, I first find the density of the joint; since the area of this is 3/2, the joint density is 2/3, everywhere.

Then the marginal for a given $Y = y$ is the length of possible X values for that y, multiplied by the density, 2/3. So for y between 0 and 1, the length is 1, so the marginal density is 2/3. For y between 1 and 2, the length is $1 - x = 1 - (y - 1) = 2 - y$, so the marginal density is (2/3)(2-y). Summing up,

$$f_Y(y) = \begin{cases} 2/3 & \text{for} \quad 0 < y < 1 \\ (2/3)(2-y) & \text{for} \quad 1 < y < 2 \\ 0 & \text{otherwise.} \end{cases}$$

Then the expected value for Y is

$$\begin{aligned} EY = \int_0^2 y f_Y(y) dy &= \int_0^1 y(2/3)dy + \int_1^2 y(2/3)(2-y)dy \\ &= y^2/3\Big|_0^1 + (2/3)(y^2 - y^3/3)\Big|_1^2 \\ &= 1/3 + (2/3)(4 - 8/3 - 1 + 1/3) = 1/3 + (2/3)(2/3) \\ &= 7/9 \end{aligned}$$

X and Y are not independent, because the shape of the support is not rectangular. Also from looking at the definition of the joint, we can tell that Y depends on X because it is uniform on the region where $Y < X + 1$.

Practice Problem 10 There are several ways to do this problem. Here are three.

1. Since they are independent, the mgf of X+Y is the product of the mgfs of X and Y, so $\phi_Z(t) = e^{4t+16t^2}$. Then take the appropriate derivatives to find the variance.
2. As above, but recognize that e^{4t+16t^2} is the mgf of a $N(4, 8)$, so the variance is 8.
3. Recognize that this is the mgf for a $N(2, 4)$, so $\text{Var}(X) = \text{Var}(Y) = 4$. Since they are independent, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) = 4 + 4 = 8$.

Practice Problem 11 The first plot has sample correlation -0.82, because it has a negative slope. Also the strength of the relationship is less than in the third plot, so we expect the absolute value of the correlation to be smaller.

The second plot has sample correlation -0.09, because it has no *linear* relationship.

The third plot has sample correlation 0.92, because it has a positive slope.

Practice Problem 12 The second histogram goes with the fourth boxplot. Both have long tails for large numbers.

The fourth histogram goes with the third boxplot. Both have no tails.

Then either note that the first histogram has generally smaller numbers than the third, or that it is slightly less spread out (that's a little hard to see though), so the the first histogram goes with the second boxplot and the third histogram goes with the first boxplot.