A simple random variable

a) The minimal sample spaces are:

\[ \Omega_x = \{0, 1\} \]
\[ \Omega_y = \{0, 1, 2\} \]

b) Let the joint PDF be defined by \( f(x_1, x_2, y) = \Pr(x_1 = x_1 \cap x_2 = x_2 \cap y = y) \). Since the sample spaces are small, the PDF can be reported by simply enumerating every possibility. Note that values such as \((x_1, x_2, y) = (0, 0, 2)\) are in the sample space and must be reported below, even though they are logically impossible.

\[
\begin{array}{cccc}
X_1 & X_2 & Y & f(X_1, X_2, Y) \\
0 & 0 & 0 & 0.25 \\
0 & 0 & 1 & 0.00 \\
0 & 0 & 2 & 0.00 \\
0 & 1 & 0 & 0.00 \\
0 & 1 & 1 & 0.25 \\
0 & 1 & 2 & 0.00 \\
1 & 0 & 0 & 0.00 \\
1 & 0 & 1 & 0.25 \\
1 & 0 & 2 & 0.00 \\
1 & 1 & 0 & 0.00 \\
1 & 1 & 1 & 0.00 \\
1 & 1 & 2 & 0.25 \\
\end{array}
\]

c) Again, we can simply enumerate:

\[
\begin{array}{cc}
Y & \Pr(y = Y | x_1 = 1) \\
0 & 0.0 \\
1 & 0.5 \\
2 & 0.5 \\
\end{array}
\]

d) The plot should look like a set of stairs, with “steps” at 0, 1, and 2.

\[
F(Y) = \Pr(y \leq Y) = \begin{cases} 
0 & \text{if } Y < 0. \\
0.25 & \text{if } 0 \leq Y < 1. \\
0.75 & \text{if } 1 \leq Y < 2. \\
1 & \text{if } 2 \leq Y. 
\end{cases}
\]
e) There are a lot of ways to figure this one out. You could calculate it directly from the probabilities, but the easiest way is to use some of our results about expected values.

\[ E(y|x_1) = E(x_1 + x_2|x_1) = x_1 + E(x_2|x_1) = x_1 + E(x_2) = x_1 + 0.5 \]

f) This one is trickier. You need to manipulate the joint and conditional probabilities directly.

\[
E(x_1|y) = 0 \cdot \Pr(x_1 = 0|y) + 1 \cdot \Pr(x_1 = 1|y) \\
= \frac{\Pr(x_1 = 1 \cap y = 0)}{\Pr(y = 0)} I(y = 0) + \frac{\Pr(x_1 = 1 \cap y = 1)}{\Pr(y = 1)} I(y = 1) + \frac{\Pr(x_1 = 1 \cap y = 2)}{\Pr(y = 2)} I(y = 2) \\
= \frac{0}{0.25} I(y = 0) + \frac{0.25}{0.5} I(y = 1) + \frac{0.25}{0.25} I(y = 2) \\
= 0.5y
\]

g) This problem can be solved by brute force, but it is easier if you notice that:

\[
\begin{align*}
var(y) &= var(x_1) + var(x_2) + 2cov(x_1, x_2) \\
&= var(x_1) + var(x_2) & \text{since } x_1 \text{ and } x_2 \text{ are independent} \\
&= 2var(x_1) & \text{since } x_1 \text{ and } x_2 \text{ are identically distributed} \\
cov(x_1, y) &= cov(x_1, x_1 + x_2) & \text{by substitution} \\
&= cov(x_1, x_1) + cov(x_1, x_2) & \text{by linearity of expectations} \\
&= cov(x_1, x_1) & \text{since } x_1 \text{ and } x_2 \text{ are independent} \\
&= var(x_1)
\end{align*}
\]

Therefore:

\[
\begin{align*}
corr(x_1, y) &= \frac{cov(x_1, y)}{\sqrt{var(x_1)var(y)}} \\
&= \frac{var(x_1)}{\sqrt{2 \cdot var(x_1)var(x_1)}} \\
&= \frac{1}{\sqrt{2}} \approx 0.71
\end{align*}
\]

2 The effects of smoking

a) The probability is just \( \frac{20,000}{30,000,000} = 0.0666666\% \).

b) We apply Bayes’ law to get:

\[
\begin{align*}
\Pr(\text{diagnosed}|\text{smoker}) &= \frac{\Pr(\text{smoker}|\text{diagnosed}) \Pr(\text{diagnosed})}{\Pr(\text{smoker})} \\
&= \frac{0.75 \times 0.0000666666}{0.2} \\
&= 0.25\%
\end{align*}
\]
c) We apply Bayes’ law to get:
\[
\Pr(\text{diagnosed}|\text{nonsmoker}) = \frac{\Pr(\text{nonsmoker}|\text{diagnosed}) \Pr(\text{diagnosed})}{\Pr(\text{nonsmoker})}
\]
\[
= \frac{0.25 \times 0.00666666}{0.8}
\]
\[
= 0.021\%
\]

d) A smoker is twelve times as likely as a nonsmoker to get lung cancer in a given year.

3 Probability theory

a) First we note that \(A_2\) can be partitioned into the pair of disjoint subsets \(A_2 \cap A_1\) and \(A_2 \cap A_1^c\). By the definition of a probability distribution:
\[
\Pr(A_2) = \Pr(A_2 \cap A_1) + \Pr(A_2 \cap A_1^c)
\]
Now, since \(A_1 \subset A_2\), we have \(A_2 \cap A_1 = A_1\). Substituting in to the equation above, we have:
\[
\Pr(A_2) = \Pr(A_1) + \Pr(A_2 \cap A_1^c)
\]
Since all probabilities are nonnegative, the result follows.

b) First, the definition of conditional expectation implies that:
\[
\Pr(A_3|A_1 \cup A_2) = \frac{\Pr(A_3 \cap (A_1 \cup A_2))}{\Pr(A_1 \cup A_2)}
\]
\[
= \frac{\Pr(A_3 \cap A_1) + \Pr(A_3 \cap A_2)}{\Pr(A_1 \cup A_2)}
\]
Since \(A_1\) and \(A_2\) are disjoint, these probabilities can be separated:
\[
\frac{\Pr((A_3 \cap A_1) \cup (A_3 \cap A_2))}{\Pr(A_1 \cup A_2)} = \frac{\Pr(A_3 \cap A_1) + \Pr(A_3 \cap A_2)}{\Pr(A_1) + \Pr(A_2)}
\]
\[
= \frac{\Pr(A_3|A_1) \Pr(A_1) + \Pr(A_3|A_2) \Pr(A_2)}{\Pr(A_1) + \Pr(A_2)}
\]

c) We have:
\[
\Pr(A_1|A_2) = \frac{\Pr(A_1 \cap A_2)}{\Pr(A_2)}
\]
\[
= \frac{\Pr(A_2|A_1) \Pr(A_1)}{\Pr(A_2)}
\]

4 Properties of expectations

a) By definition we have:
\[
\text{var}(X) = E \left[ (X - E(X))^2 \right]
\]
\[
= E \left[ X^2 - 2XE(X)E(X)^2 \right]
\]
\[
= E(X^2) - E(X)^2
\]
b) We have:
\[
\text{cov}(aX + bY, cX + dY) = E \left[ ((aX + bY) - E(aX + bY)) \cdot ((cX + dY) - E(cX + dY)) \right] \\
= E \left[ (a(X - E(X)) + b(Y - E(Y))) \cdot (c(X - E(X)) + d(Y - E(Y))) \right] \\
= ac \text{ var}(X) + bd \text{ var}(Y) + (ad + bc) \text{ cov}(X, Y)
\]

c) First, we note that
\[
E(g(X)Y|X) = g(X)E(Y|X) = g(X) \cdot 0 = 0
\]
Taking expectations of both sides, and applying the law of iterated expectations we have:
\[
E(g(X)Y) = E(E(g(X)Y|X)) = E(0)
\]

5 Basic data manipulation in R

This code will work:
```r
rbern <- function(n,k,p) {
  x1 <- runif(n*k)
  x2 <- as.integer(x1 < p)
  matrix(x2, nrow=n)
}
x <- rbern(3,5,0.75)
print(x)
apply(x,2,mean)
```

It wouldn’t be following the instructions, but the `rbern` function could be written more briefly as:
```r
rbern <- function(n,k,p) matrix(as.integer(runif(n*k)<p),nrow=n)
```