# Probability distributions
## Lectures, questions and solutions

Nicky van Foreest and Ruben van Beesten

February 10, 2021

CONTENTS

## 1   LECTURE 1

**Ex 1.1.** Consider 12 football players on a football field. Eleven of them are players of FC Barcelone, the other one is an arbiter. We select a random player, uniform. This player must take a penalty. The probability that a player of Barcelone scores is 70%, for the arbiter it is 50%. Let $P \in \{A, B\}$ be r.v that corresponds to the selected player, and $S \in \{0, 1\}$ be the score.

1. What is the PMF? In other words, determine $P\{P = B, S = 1\}$ and so on for all possibilities.

2. What is $P\{S = 1\}$? What is $P\{P = B\}$?

3. Show that $S$ and $P$ are dependent.

An insurance company receives on a certain day two claims $X, Y \geq 0$. We will find the PMF of the loss $Z = X + Y$ under different assumptions.

The joint CDF $F_{X,Y}$ and joint PMF $p_{X,Y}$ are assumed known.

**Ex 1.2.** Why is it not interesting to consider the case $\{X = 0, Y = 0\}$?

**Ex 1.3.** Find an expression for the PMF of $Z = X + Y$.

Suppose $p_{X,Y}(i,j) = c\, I_{i=j} I_{1 \leq i \leq 4}$.

**Ex 1.4.** What is $c$?

**Ex 1.5.** What is $F_X(i)$? What is $F_Y(j)$?

**Ex 1.6.** Are $X$ and $Y$ dependent? If so, why, because $1 = F_{X,Y}(4, 4) = F_X(4)F_Y(4)$?

**Ex 1.7.** What is $P\{Z = k\}$?

**Ex 1.8.** What is $V[Z]$?

Now take $X, Y$ iid $\sim \text{Unif}(\{1, 2, 3, 4\})$ (so now no longer $p_{X,Y}(i,j) = I_{i=j} I_{1 \leq i \leq 4}$).

**Ex 1.9.** What is $P\{Z = 4\}$?

**Remark 1.1.** We can make lots of variations on this theme.

1. Let $X \in \{1, 2, 3\}$ and $Y \in \{1, 2, 3, 4\}$.

2. Take $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$. (Use the chicken-egg story)

3. We can make $X$ and $Y$ such that they are (both) continuous, i.e., have densities. The conceptual ideas[1] don't change much, except that the summations become integrals.

4. Why do people often/sometimes (?) model the claim sizes as iid $\sim \text{Norm}(\mu, \sigma^2)$? There is a slight problem with this model (can real claim sizes be negative?), but what is the way out?

5. The example is more versatile than you might think. Here is another interpretation.

   A supermarket has 5 packets of rice on the shelf. Two customers buy rice, with amounts $X$ and $Y$. What is the probability of a lost sale, i.e., $P\{X + Y > 5\}$? What is the expected amount lost, i.e., $E[\max X + Y - 5, 0]$?

   Here is yet another. Two patients arrive in to the first aid of a hospital. They need $X$ and $Y$ amounts of service, and there is one doctor. When both patients arrive at 2 pm, what is the probability that the doctor has work in overtime (after 5 pm), i.e., $P\{X + Y > 5 - 2\}$?

---

[1] Unless you start digging deeper. Then things change drastically, but we skip this technical stuff.

## 2   LECTURE 2

Read the problems of `memoryless\_excursions.pdf`. All the problems in that document relate to topics discussed in Sections BH.7.1 and BH.7.2, and quite a lot of topics you have seen in the previous course on probability theory.

## 3  LECTURE 3

**Ex 3.1.** We ask a married woman on the street her height $X$. What does this tell us about the height $Y$ of her spouse? We suspect that taller/smaller people choose taller/smaller partners, so, given $X$, a simple estimator $\hat{Y}$ of $Y$ is given by

$$\hat{Y} = aX + b.$$

(What is the sign of $a$ if taller people tend to choose taller people as spouse?) But how to determine $a$ and $b$? A common method is to find $a$ and $b$ such that the function

$$f(a,b) = \mathsf{E}\left[(Y - \hat{Y})^2\right]$$

is minimized. Show that the optimal values are such that

$$\hat{Y} = \mathsf{E}[Y] + \rho \frac{\sigma_Y}{\sigma_X}(X - \mathsf{E}[X]),$$

where $\rho$ is the correlation between $X$ and $Y$ and where $\sigma_X$ and $\sigma_Y$ are the standard deviations of $X$ and $Y$ respectively.

**Ex 3.2.** Using scaling laws often can help to find errors. For instance, the prediction $\hat{Y}$ should not change whether we measure the height in meters or centimetres. In view of this, explain that

$$\hat{Y} = \mathsf{E}[Y] + \rho \frac{\mathsf{V}[Y]}{\sigma_X}(X - \mathsf{E}[X])$$

must be wrong.

**Ex 3.3.** $N$ people throw their hat in a box. After shuffling, each of them takes out a hat at random. How many people do you expect to take out their own hat (i.e., the hat they put in the box); what is the variance? In BH.7.46 you have to solve this analytically. In the exercise here you have to write a simulator for compute the expectation and variance.

INDICATORS ARE GREAT    functions, and I suspect you underestimated the importance of these functions. They help to keep your formulas clean. You can use them in computer code as logical conditions, or to help counting relevant events, something you need when numerically estimating multi-D integrals for machine learning for instance. And, even though I(=NvF) often prefer to use figures over algebra to understand something, when it comes to integration (and reversing the sequence of integration in multiple integrals) I find indicators easier to use.

Moreover, you should know that in fact, *expectation* is the fundamental concept in probability theory, and the *probability of an event is defined* as

$$\mathsf{P}\{A\} := \mathsf{E}[I_A]. \tag{3.1}$$

Thus, the fundamental bridge is actually an application of LOTUS to indicator functions. REREAD BH.4.4!

Here are some more examples

**Ex 3.4.** What is $\int_{-\infty}^{\infty} I_{0 \leq x \leq 3}\, \mathrm{d}x$?

**Ex 3.5.** What is

$$\int x\, I_{0 \le x \le 4}\, \mathrm{d}x? \tag{3.2}$$

When we do an integral over a 2D surface we can first integrate over the $x$ and then over the $y$, or the other way around, whatever is the most convenient. (There are conditions about how to handle multi-D integral, but for this course these are irrelevant.)

**Ex 3.6.** What is

$$\iint xy\, I_{0 \le x \le 3}\, I_{0 \le y \le 4}\, \mathrm{d}x\, \mathrm{d}y? \tag{3.3}$$

**Ex 3.7.** What is

$$\iint I_{0 \le x \le 3}\, I_{0 \le y \le 4}\, I_{x \le y}\, \mathrm{d}x\, \mathrm{d}y? \tag{3.4}$$

## 4   LECTURE 4

**Ex 4.1.** BH.7.65 Let $(X_1,\ldots,X_k)$ be Multinomial with parameters $n$ and $(p_1,\ldots,p_k)$. Use indicator r.v.s to show that $\text{Cov}(X_i,X_j) = -np_ip_j$ for $i \neq j$.

**Ex 4.2.** Suppose $(X,Y)$ are bivariate normal distributed with mean vector $\mu = (\mu_X,\mu_Y) = (0,0)$, standard deviations $\sigma_X = \sigma_Y = 1$ and correlation $\rho_{XY}$ between $X$ and $Y$. Specify the joint pdf of $X$ and $X+Y$.

The following exercises will show how probability theory can be used in finance. We will look at the tradeoff between risk and return in a financial portfolio.

John is an investor who has \$10,000 to invest. There are three stocks he can choose from. The returns on investment $(A,B,C)$ of these three stocks over the following year (in terms of percentages) follow a multinomial distribution. The expected returns on investment are $\mu_A = 7.5\%$, $\mu_B = 10\%$, $\mu_C = 20\%$. The corresponding standard deviations are $\sigma_A = 7\%$, $\sigma_B = 12\%$ and $\sigma_C = 17\%$. Note that risk (measured in standard deviation) increases with expected return. The correlation coefficients between the different returns are $\rho_{AB} = 0.7$, $\rho_{AC} = -0.8$, $\rho_{BC} = -0.3$.

**Ex 4.3.** Suppose the investor decides to invest \$2,000 in stock A, \$4,000 in stock B, \$2,000 in stock C and to put the remaining \$2,000 in a savings account with a zero interest rate. What the expected value of his portfolio after a year?

**Ex 4.4.** What is the standard deviation of the value of the portfolio in a year?

**Ex 4.5.** John does not like losing money. What is his probability of having made a net loss after a year?

John has a friend named Mary, who is a first-year EOR student. She has never invested money herself, but she is paying close attention during the course Probability Distributions. She tells her friend: "John, your investment plan does not make a lot of sense. You can easily get a higher expected return at a lower level of risk!"

**Ex 4.6.** Show that Mary is right. That is, make a portfolio with a higher expected return, but with a lower standard deviation.
*Hint: Make use of the **negative correlation** between C and the other two stocks!*

## 5  SOLUTIONS

**s.1.1.** Here is the joint PMF:

$$\mathsf{P}\{P = A, S = 1\} = \frac{1}{12}0.5 \qquad\qquad \mathsf{P}\{P = A, S = 0\} = \frac{1}{12}0.5 \tag{5.1}$$

$$\mathsf{P}\{P = B, S = 1\} = \frac{11}{12}0.7 \qquad\qquad \mathsf{P}\{P = B, S = 0\} = \frac{11}{12}0.3. \tag{5.2}$$

Now the marginal PMFs

$$\mathsf{P}\{S = 1\} = \mathsf{P}\{P = A, S = 1\} + \mathsf{P}\{P = B, S = 1\} = 0.042 + 0.64 = 0.683 = 1 - \mathsf{P}\{S = 0\}$$

$$\mathsf{P}\{P = B\} = \frac{11}{12} = 1 - \mathsf{P}\{P = A\}.$$

For independence we take the definition. In general, for all outcomes $x, y$ we must have that $\mathsf{P}\{X = x, Y = y\} = \mathsf{P}\{X = x\}\,\mathsf{P}\{Y = y\}$. For our present example, let's check for a particular outcome:

$$\mathsf{P}\{P = B, S = 1\} = \frac{11}{12} \cdot 0.7 \neq \mathsf{P}\{P = B\}\,\mathsf{P}\{S = 1\} = \frac{11}{12} \cdot 0.683$$

The joint PMF is obviously not the same as the product of the marginals, which implies that $P$ and $S$ are not independent.

**s.1.2.** When the claim sizes are 0, then the insurance company does not receive a claim.

**s.1.3.** By the fundamental bridge,

$$\mathsf{P}\{Z = k\} = \sum_{i,j} I_{i+j=k}\, p_{X,Y}(i,j) \tag{5.3}$$

$$= \sum_{i,j} I_{i,j \geq 0}\, I_{j=k-i}\, p_{X,Y}(i,j) \tag{5.4}$$

$$= \sum_{i=0}^{k} p_{X,Y}(i, k - i). \tag{5.5}$$

**s.1.4.** $c = 1/4$ because there are just four possible values for $i$ and $j$.

**s.1.5.** Use marginalization:

$$F_X(k) = F_{X,Y}(k, \infty) = \sum_{i \leq k} \sum_{j} p_{X,Y}(i,j) \tag{5.6}$$

$$= \frac{1}{4} \sum_{i \leq k} \sum_{j} I_{i=j}\, I_{1 \leq i \leq 4} \tag{5.7}$$

$$= \frac{1}{4} \sum_{i \leq k} I_{1 \leq i \leq 4} \tag{5.8}$$

$$= k/4, \tag{5.9}$$

$$F_Y(j) = j/4. \tag{5.10}$$

**s.1.6.** The equality in the question must hold for all $i, j$, not only for $i = j = 4$. If you take $i = j = 1$, you'll see immediately that $F_{X,Y}(1,1) \neq F_X(1)F_Y(1)$:

$$\frac{1}{4} = F_{X,Y}(1,1) \neq F_X(1)F_Y(1) = \frac{1}{4}\frac{1}{4}. \tag{5.11}$$

**s.1.7.** $P\{Z = 2\} = P\{X = 1, Y = 1\} = 1/4 = P\{Z = 4\}$, etc. $P\{Z = k\} = 0$ for $k \notin \{2, 4, 6, 8\}$.

**s.1.8.** Here is one approach

$$V[Z] = E\left[Z^2\right] - (E[Z])^2 \tag{5.12}$$

$$E\left[Z^2\right] = E\left[(X + Y)^2\right] = E\left[X^2\right] + 2E[XY] + E\left[Y^2\right] \tag{5.13}$$

$$(EZ)^2 = (E[X] + E[Y])^2 \tag{5.14}$$

$$= (E[X])^2 + 2E[X]E[Y] + (E[Y])^2 \tag{5.15}$$

$$\implies \tag{5.16}$$

$$V[Z] = E\left[Z^2\right] - (E[Z])^2 \tag{5.17}$$

$$= V[X] + V[Y] + 2(E[XY] - (E[X]E[Y])) \tag{5.18}$$

$$E[XY] = \sum_{ij} ij p_{X,Y}(i,j) = \frac{1}{4}(1 + 4 + 9 + 16) = \dots \tag{5.19}$$

$$E\left[X^2\right] = \dots \tag{5.20}$$

The numbers are for you to compute.

**s.1.9.**

$$P\{Z = 4\} = \sum_{i,j} I_{i+j=4} p_{X,Y}(i,j) \tag{5.21}$$

$$= \sum_{i=1}^{4} \sum_{j=1}^{4} I_{j=4-i} \frac{1}{16} \tag{5.22}$$

$$= \sum_{i=1}^{3} \frac{1}{16} \tag{5.23}$$

$$= \frac{3}{16}. \tag{5.24}$$

**s.3.1.** We take the partial derivatives of $f$ with respect to $a$ and $b$, and solve for $a$ and $b$. In the derivation, we use that

$$\rho = \frac{\text{Cov}[X,Y]}{\sqrt{V[X]V[Y]}} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y} \implies \rho \frac{\sigma_Y}{\sigma_X} = \frac{\text{Cov}[X,Y]}{V[X]}. \tag{5.25}$$

Hence,

$$
\begin{aligned}
f(a,b) &= \mathsf{E}\big[(Y - \hat{Y})^2\big] \\
&= \mathsf{E}\big[(Y - aX - b)^2\big] \\
&= \mathsf{E}\big[Y^2\big] - 2a\,\mathsf{E}[YX] - 2b\,\mathsf{E}[Y] + a^2\,\mathsf{E}\big[X^2\big] + 2ab\,\mathsf{E}[X] + b^2 \\
\partial_a f &= -2\,\mathsf{E}[YX] + 2a\,\mathsf{E}\big[X^2\big] + 2b\,\mathsf{E}[X] = 0 \\
&\implies a\,\mathsf{E}\big[X^2\big] = \mathsf{E}[YX] - b\,\mathsf{E}[X] \\
\partial_b f &= -2\,\mathsf{E}[Y] + 2a\,\mathsf{E}[X] + 2b = 0 \\
&\implies b = \mathsf{E}[Y] - a\,\mathsf{E}[X] \\
a\,\mathsf{E}\big[X^2\big] &= \mathsf{E}[YX] - \mathsf{E}[X](\mathsf{E}[Y] - a\,\mathsf{E}[X]) \\
&\implies a(\mathsf{E}\big[X^2\big] - \mathsf{E}[X]\,\mathsf{E}[X]) = \mathsf{E}[YX] - \mathsf{E}[X]\,\mathsf{E}[Y] \\
&\implies a = \frac{\mathsf{Cov}[X,Y]}{\mathsf{V}[X]} = \rho\frac{\sigma_Y}{\sigma_X} \\
b &= \mathsf{E}[Y] - \rho\frac{\sigma_Y}{\sigma_X}\mathsf{E}[X] \\
\hat{Y} &= aX + b \\
&= \rho\frac{\sigma_Y}{\sigma_X}X + \mathsf{E}[Y] - \rho\frac{\sigma_Y}{\sigma_X}\mathsf{E}[X] \\
&= \mathsf{E}[Y] + \rho\frac{\sigma_Y}{\sigma_X}(X - \mathsf{E}[X]).
\end{aligned}
$$

What a neat formula! Memorize the derivation, at least the structure. You'll come across many more optimization problems.

What if $\rho = 0$?

**s.3.2.** If we measure $X$ in centimetres instead of metres, then $X$, $\mathsf{E}[X]$ and $\sigma_X$ are all multiplied by 100, and the prediction $\hat{Y}$ should also be expressed in centimetres But $\mathsf{V}[Y]$ scales as length squared. This messes up the units.

**s.3.3.** Let us first do one run.

```python
import numpy as np

np.random.seed(3)

N = 4
X = np.arange(N)
np.random.shuffle(X)
print(X)
print(np.arange(N))
print((X == np.arange(N)))
print((X == np.arange(N)).sum())
```

Here are the results of the print statements: $X = [3102]$. The matches are [False True False False]; we see that $X[1] = 1$ (recall, python arrays start at index 0, not at 1, so $X[1]$ is the second element of $X$, not the first), so that the second person picks his own hat. The number of matches is therefore 1 for this simulation.

Now put the people to work, and let them pick hats for 50 times.

```python
import numpy as np

np.random.seed(3)

num_samples = 50
N = 5

res = np.zeros(num_samples)
for i in range(num_samples):
    X = np.arange(N)
    np.random.shuffle(X)
    res[i] = (X == np.arange(N)).sum()

print(res.mean(), res.var())
```

Here is the number of matches for each round: [0. 1. 1. 0. 1. 0. 1. 0. 1. 1. 1. 2. 2. 1. 0. 1. 1. 1. 0. 2. 0. 1. 2. 2. 0. 0. 0. 1. 0. 1. 3. 1. 1. 2. 3. 0. 1. 0. 3. 1. 2. 0. 2. 0. 1. 0. 3. 0. 1. 0.] The mean and variance are as follows: $E[X] = 0.96$ and $V[X] = 0.8384$.

For your convenience, here's the R code

```r
# set seed such that results can be recreated
set.seed(42)

# number simulations and people
numSamples <- 50
N <- 5

# initialize empty result vector
res <- c()

# for loop to simulate repeatedly
for (i in 1:numSamples) {

  # shuffle the N hats
  x <- sample(1:N)

  # number of people picking own hat (element by element the vectors x and
  # 1:N are compared, which yields a vector of TRUE and FALSE, TRUE = 1 and
  # FALSE = 0)
  correctPicks <- sum(x == 1:N)

  # append the result vector by the result of the current simulation
  res <- append(res, correctPicks)
}

# printing of observed mean and variance
```

```
27   print(mean(res))
28   print(var(res))
```

**s.3.4.**

$$\int_{-\infty}^{\infty} I_{0 \le x \le 3} \, dx = \int_{0}^{3} dx = 3.$$

**s.3.5.**

$$\int x \, I_{0 \le x \le 4} \, dx = \int_{0}^{4} x \, dx = \ldots$$

**s.3.6.**

$$\iint xy \, I_{0 \le x \le 3} \, I_{0 \le y \le 4} \, dx \, dy = \int_{0}^{3} x \int_{0}^{4} y \, dy \, dx$$

$$= \int_{0}^{3} x \frac{y^2}{2} \Big|_{0}^{4} \, dx$$

$$= \int_{0}^{3} x \, dx \, 8 = \ldots$$

**s.3.7.** Two solutions. First we integrate over $y$.

$$\iint I_{0 \le x \le 3} \, I_{0 \le y \le 4} \, I_{x \le y} \, dx \, dy = \int I_{0 \le x \le 3} \int I_{0 \le y \le 4} \, I_{x \le y} \, dy \, dx \tag{5.26}$$

$$= \int I_{0 \le x \le 3} \int I_{\max\{x,0\} \le y \le 4} \, dy \, dx \tag{5.27}$$

$$= \int_{0}^{3} \int_{\max\{x,0\}}^{4} dy \, dx \tag{5.28}$$

$$= \int_{0}^{3} y \Big|_{\max\{x,0\}}^{4} dx \tag{5.29}$$

$$= \int_{0}^{3} (4 - \max\{x,0\}) \, dx \tag{5.30}$$

$$= 12 - \int_{0}^{3} \max\{x,0\} \, dx \tag{5.31}$$

$$= 12 - \int_{0}^{3} x \, dx \tag{5.32}$$

$$= 12 - 9/2. \tag{5.33}$$

Let's now instead first integrate over $x$.

$$\iint I_{0 \leq x \leq 3} I_{0 \leq y \leq 4} I_{x \leq y} \, \mathrm{d}x \, \mathrm{d}y = \int I_{0 \leq y \leq 4} \int I_{0 \leq x \leq 3} I_{x \leq y} \, \mathrm{d}x \, \mathrm{d}y \tag{5.34}$$

$$= \int_0^4 \int I_{0 \leq x \leq \min\{3, y\}} \, \mathrm{d}x \, \mathrm{d}y \tag{5.35}$$

$$= \int_0^4 \int_0^{\min\{3, y\}} \mathrm{d}x \, \mathrm{d}y \tag{5.36}$$

$$= \int_0^4 \min\{3, y\} \, \mathrm{d}y \tag{5.37}$$

$$= \int_0^3 \min\{3, y\} \, \mathrm{d}y + \int_3^4 \min\{3, y\} \, \mathrm{d}y \tag{5.38}$$

$$= \int_0^3 y \, \mathrm{d}y + \int_3^4 3 \, \mathrm{d}y \tag{5.39}$$

$$= 9/2 + 3. \tag{5.40}$$

**s.4.1.** See solution manual.

**s.4.2.** Define $V := X$ and $W := X + Y$. Observe that for any $t_V, t_W$, we have

$$t_V V + t_W W = t_V X + t_W (X + Y) \tag{5.41}$$

$$= (t_V + t_W) X + t_W Y. \tag{5.42}$$

Hence, any linear combination of $V$ and $W$ is a linear combination of $X$ and $Y$. Since $(X, Y)$ is bivariate normal, every linear combination of $X$ and $Y$ is normally distributed. Hence, every linear combination of $V$ and $W$ is normally distributed. Hence, by definition, $(V, W)$ is bivariate normally distributed.

We need to compute the mean vector and covariance matrix of $(V, W)$. We have

$$\mu_V = \mathsf{E}[V] = \mathsf{E}[X] = \mu_X = 0, \tag{5.43}$$

and

$$\mu_W = \mathsf{E}[W] = \mathsf{E}[X + Y] = \mu_X + \mu_Y = 0. \tag{5.44}$$

Next, we have

$$\mathsf{V}[V] = \mathsf{V}[X] = \sigma_X^2 = 1, \tag{5.45}$$

and

$$\mathsf{V}[W] = \mathsf{V}[X + Y] = \mathsf{V}[X] + \mathsf{V}[Y] + 2\mathrm{Cov}(X, Y) \tag{5.46}$$

$$= 1 + 1 + 2\rho_{XY} \sigma_X \sigma_Y = 2(1 + \rho_{XY}). \tag{5.47}$$

Finally,

$$\mathrm{Cov}(V, W) = \mathrm{Cov}(X, X + Y) = \mathrm{Cov}(X, X) + \mathrm{Cov}(X, Y) \tag{5.48}$$

$$= \sigma_X^2 + \rho_{XY} \sigma_X \sigma_Y = 1 + \rho_{XY}, \tag{5.49}$$

and hence,

$$\rho_{VW} := \mathrm{Corr}(V, W) = \frac{\mathrm{Cov}(V, W)}{\sqrt{\mathsf{V}[V]\,\mathsf{V}[W]}} \tag{5.50}$$

$$= \frac{1 + \rho_{XY}}{\sqrt{1 \cdot 2(1 + \rho_{XY})}} \tag{5.51}$$

$$= \sqrt{\frac{1 + \rho_{XY}}{2}}. \tag{5.52}$$

We have now specified all parameters of the bivariate normal distribution. This yields the following joint pdf:

$$f_{V,W}(v, w) = \frac{1}{2\pi\sigma_V\sigma_W\tau_{VW}}\exp\left(-\frac{1}{2\tau_{VW}^2}\left(\left(\frac{v}{\sigma_V}\right)^2 + \left(\frac{w}{\sigma_W}\right)^2 - 2\frac{\rho_{VW}}{\sigma_V\sigma_W}vw\right)\right), \tag{5.53}$$

where $\tau_{VW} := \sqrt{1 - \rho_{VW}^2} = \sqrt{1 - \frac{1+\rho_{XY}}{2}} = \sqrt{\frac{1-\rho_{XY}}{2}}$ and $\sigma_V = \sqrt{\mathsf{V}[V]} = 1$ and $\sigma_W = \sqrt{\mathsf{V}[W]} = \sqrt{2(1 + \rho_{XY})}$. Hence,

$$f_{V,W}(v, w) = \frac{1}{2\pi\sqrt{1 - (\rho_{XY})^2}}\,\mathsf{E}\left[\exp\left(-\frac{1}{1 - \rho_{XY}}\left(v^2 + \frac{w^2}{2(1 + \rho_{XY})} - \frac{1}{2}vw\right)\right)\right]. \tag{5.54}$$

**s.4.3.** Let $X$ denote the value of the portfolio after a year in thousands of dollars. Then,

$$X := 2(1 + A) + 4(1 + B) + 2(1 + C) + 2 \tag{5.55}$$

$$= 10 + 2A + 4B + 2C. \tag{5.56}$$

Then,

$$\mathsf{E}[X] = \mathsf{E}[10 + 2A + 4B + 2C] \tag{5.57}$$

$$= 10 + 2\,\mathsf{E}[A] + 4\,\mathsf{E}[B] + 2\,\mathsf{E}[C] \tag{5.58}$$

$$= 10 + 2 \cdot 0.075 + 4 \cdot 0.1 + 2 \cdot 0.2 \tag{5.59}$$

$$= 10 + 0.15 + 0.4 + 0.4 \tag{5.60}$$

$$= 10.95 \tag{5.61}$$

**s.4.4.** We have

$$\mathsf{V}[X] = \mathsf{V}[10 + 2A + 4B + 2C] \tag{5.62}$$

$$= \mathsf{V}[2A] + \mathsf{V}[4B] + \mathsf{V}[2C] \tag{5.63}$$

$$\quad + 2\Big(\mathrm{Cov}(2A, 4B) + \mathrm{Cov}(2A, 2C) + \mathrm{Cov}(4B, 2C)\Big) \tag{5.64}$$

$$= 4\mathsf{V}[A] + 16\mathsf{V}[B] + 4\mathsf{V}[C] \tag{5.65}$$

$$\quad + 2\Big(8\mathrm{Cov}(A, B) + 4\mathrm{Cov}(A, C) + 8\mathrm{Cov}(B, C)\Big) \tag{5.66}$$

$$= 4\sigma_A^2 + 16\sigma_B^2 + 4\sigma_C^2 \tag{5.67}$$

$$\quad + 2\Big(8\rho_{AB}\sigma_A\sigma_B + 4\rho_{AC}\sigma_A\sigma_C + 8\rho_{BC}\sigma_B\sigma_C\Big) \tag{5.68}$$

$$= 4(0.07)^2 + 16(0.12)^2 + 4(0.17)^2 \tag{5.69}$$

$$\quad + 2\Big(8(0.7)(0.07)(0.12) + 4(-0.8)(0.07)(0.17) + 8(-0.3)(0.12)(0.17)\Big) \tag{5.70}$$

$$= 0.2856. \tag{5.71}$$

So

$$\sigma_X = \sqrt{0.2856} = 0.5344. \tag{5.72}$$

So $X$ has a standard deviation of \$534.

**s.4.5.** We need to compute the probability $\mathsf{P}\{X \le 10\}$. We have

$$\mathsf{P}\{X \le 10\} = \mathsf{P}\left\{X - \mu_X \le 10 - 10.95\right\} \tag{5.73}$$

$$= \mathsf{P}\left\{\frac{X - \mu_X}{\sigma_X} \le \frac{10 - 10.95}{0.5344}\right\} \tag{5.74}$$

$$= \mathsf{P}\left\{Z \le \frac{10 - 10.95}{0.5344}\right\} \tag{5.75}$$

$$= \mathsf{P}\{Z \le -1.9625\} \tag{5.76}$$

$$= 0.0377. \tag{5.77}$$

So John has a probability of 3.77% of losing money with his investment.

**s.4.6.** Observe that $C$ has the highest expected return *and* it is negatively correlated with the other two stocks. We will use these facts to our advantage.

Starting out with portfolio $X$, we construct a portfolio $Y$ by splitting the investment in stock B in two halves, which we add to our investments in stock A and C. Since the average expected return of A and C is higher than that of B, we must have that $\mathsf{E}[Y] > \mathsf{E}[X]$. Moreover, the fact that A and C are negatively correlated will mitigate the level of risk. If one stock goes up, we expect the other to go down, so the stocks cancel out each others variability. This is the idea behind the investment principle of *diversification*.

Mathematically, we define

$$Y := 4(1 + A) + 4(1 + C) + 2 \tag{5.78}$$

$$= 10 + 4A + 4C. \tag{5.79}$$

Then,

$$\mathsf{E}[Y] = \mathsf{E}[10 + 4A + 4C] \tag{5.80}$$

$$= 10 + 4\,\mathsf{E}[A] + 4\,\mathsf{E}[C] \tag{5.81}$$

$$= 10 + 4(0.075) + 4(0.20) \tag{5.82}$$

$$= 11.1 \tag{5.83}$$

Moreover,

$$\mathsf{V}[Y] = \mathsf{V}[10 + 4A + 4C] \tag{5.84}$$

$$= \mathsf{V}[4A] + \mathsf{V}[4C] + 2\mathrm{Cov}(4A, 4C) \tag{5.85}$$

$$= 4^2\,\mathsf{V}[A] + 4^2\,\mathsf{V}[C] + 2 \cdot 4 \cdot 4 \cdot \mathrm{Cov}(A, C) \tag{5.86}$$

$$= 16(.07)^2 + 16(.17)^2 + 32(-.8)(.07)(.17) \tag{5.87}$$

$$= 0.23616, \tag{5.88}$$

which corresponds to a standard deviation of

$$\sigma_Y = \sqrt{\mathsf{V}[Y]} = \sqrt{0.23616} = 0.4860 \tag{5.89}$$

So indeed, $\mathsf{E}[Y] > \mathsf{E}[X]$, while $\sigma_Y < \sigma_X$. Clearly, portfolio $Y$ is more desirable.