Solutions to selected exercises from Chapter 3 of *Wasserman — All of Statistics*

(1) Suppose you halve your money m times out of n. There are $\binom{n}{m}$ ways to do this, each with probability $(1/2)^n$. Your fortune after n trials is $(1/2)^m 2^n c$. We get the expected value of your fortune after n trials by summing over m:

$$\sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^m 2^{n-m} c = \sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{2}\right)^{2m} c.$$

We can recognize this as

$$\sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{4}\right)^m (1)^{n-m} = c\left(1+\frac{1}{4}\right)^n = c\left(\frac{5}{4}\right)^n.$$

(3) We suppose the X_i are independent. For $x \in [0,1]$, the CDF of Y satisfies

$$F_Y(x) = \mathbb{P}(X_i \leq x \text{ for all } i) = x^n$$

and $f_Y(x) = nx^{n-1}$ on $[0,1]$ and 0 elsewhere.
Thus

$$\mathbb{E}(Y) = \int_0^1 x(nx^{n-1})dx = \frac{n}{n+1}.$$

(4) Let $Y_1, \ldots, Y_n$ be independent Bernoulli(p) random variables with $Y_i = 0$ if the particle moves to the right on the i^{th} step and $Y_i = 1$ otherwise. Then

$$X_n = \sum_{i=1}^{n}(1-2Y_i).$$

We have $\mathbb{E}(X_n) = \sum_{i=0}^{n} \mathbb{E}(1-2Y_i)$ and $\mathbb{E}(1-2Y_i) = (1-p)(1) + p(-1) = 1-2p$ so $\mathbb{E}(X_n) = n(1-2p)$.
By independence, the variance is $\mathbb{V}(X_n) = \sum_{i=0}^{n} \mathbb{V}(1-2Y_i)$. We have $(1-2Y_i)^2 = 1$ always so $\mathbb{E}(((1-2Y_i)^2) = 1$. So

$$\mathbb{V}(1-2Y_i) = \mathbb{E}((1-2Y_i)^2) - \mathbb{E}(1-2Y_i)^2 = 1-(1-2p)^2 = 4p-4p^2 \text{ and } \mathbb{V}(X_n) = 4np(1-p).$$

(5) The sample space is $\{H, TH, TTH, TTTH, \ldots\}$ with the i^{th} entry occurring with probability $1/2^i$. The expected value is thus

$$\frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(3) + \ldots = \sum_{i=0}^{\infty} \frac{i}{2^i}.$$

Recall that the Taylor series of $1/(1-x^2)$ is $\sum i x^{i-1}$. So the expected value is

$$\frac{1}{2}\sum_{i=0}^{\infty} \frac{i}{2^{i-1}} = \frac{1}{2}\frac{1}{(1-\frac{1}{2})^2} = 2.$$

(9) See the Jupyter notebook 9.ipynb.

(10) By the rule of the lazy statistician,

$$\mathbb{E}(Y) = \frac{1}{\sqrt{2\pi}} \int e^x e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int e^{-(x-1)^2/2+1/2} dx = \sqrt{e} \int \frac{1}{\sqrt{2\pi}} e^{-(x-1)^2/2} dx.$$

The integral $\int \frac{1}{\sqrt{2\pi}} e^{-(x-1)^2/2} dx$ is just the total integral of the PDF of a $N(1,1)$ random variable so it's equal to 1. Finally then $\mathbb{E}(Y) = \sqrt{e}$.
To compute the variance we consider $Y^2 = e^{2X}$. We have

$$\mathbb{E}(Y^2) = \int \frac{1}{\sqrt{2\pi}} e^{1/2(4x-x^2)} dx = \int \frac{1}{\sqrt{2\pi}} e^{-1/2(2-x)^2+2} dx = e^2 \int \frac{1}{\sqrt{2\pi}} e^{-1/2(2-x)^2} dx = e^2.$$

Finally then

$$\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = e^2 - e.$$

(11) (a) First

$$\mathbb{E}(X_n) = \sum \mathbb{E}(X_i) = 0.$$

We have $\mathbb{V}(X_i) = \frac{1}{2}(1-0)^2 + \frac{1}{2}(-1-0)^2 = 1$ so

$$\mathbb{V}(X_n) = \sum \mathbb{V}(X_i) = n.$$

(b) See the Jupyter notebook 11.ipynb.

(15) Set $Z = r(X,Y) = 2X - 3Y$. We have $\mathbb{V}(2X - 3Y + 8) = \mathbb{V}(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2$. We'll calculate these two terms separately.

$$\mathbb{E}(Z) = \int_0^2 \int_0^1 (2x - 3y) \frac{1}{3}(x+y) dx dy = -\frac{23}{9}.$$

And $Z^2 = 4X^2 - 12XY + 9Y^2$ so

$$\mathbb{E}(Z^2) = \int_0^2 \int_0^1 (4x^2 - 12xy + 9y^2) \frac{1}{3}(x+y) dx dy = \frac{86}{9}.$$

So

$$\mathbb{V}(Z) = \frac{86}{9} - \left(\frac{23}{9}\right)^2 = \frac{245}{81}.$$

(21) By hypothesis,

$$x = \mathbb{E}(Y|X = x) = \int y \frac{f_{X,Y}(x,y)}{f_X(x)} dy$$

i.e. $\int y f_{X,Y}(x,y) dy = x f_X(x)$.
Letting μ_X, μ_Y be the expectations, we have

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X - \mu_X)(Y - \mu_Y)) = \int \int (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) dx dy$$

and this is equal to

$$\int (x - \mu_X) \left(\int (y - \mu_Y) f_{X,Y}(x,y) dy \right) dx = \int (x - \mu_X) \left(x f_X(x) - \mu_Y f_X(x) \right) dx$$

using that $\int y f_{X,Y}(x,y)dy = x f_X(x)$.

Now note by the Rule of Iterated Expectations that

$$\mu_Y = \mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(X) = \mu_X.$$

plugging this into the equality from the last paragraph gives

$$\mathrm{Cov}(X,Y) = \int (x-\mu_X)(x-\mu_Y) f_X(x)dx = \int (x-\mu_X)^2 f_X(x)dx = \mathbb{V}(X).$$

(22) (a) Y and Z are not independent. For $\mathbb{P}(Y=1, Z=1) = b-a$ whereas $\mathbb{P}(Y=1)\mathbb{P}(Z=1) = b(1-a) > b-a$.

(b) $\mathbb{E}(Y|Z=z) = \sum y f_{Y|Z}(y|z)$ and $f_{Y|Z}(y|z) = f_{Y,Z}(y,z)/f_Z(z)$. The distribution $f_{Y,Z}$ is given by the following table:

y/z	0	1
0	0	$1-b$
1	a	$b-a$

We have $f_Z(0) = a$ and $f_Z(1) = 1-a$. So $f_{Y|Z}$ is given by the following table:

y/z	0	1
0	0	$\frac{1-b}{1-a}$
1	1	$\frac{b-a}{1-a}$

Thus,

$$\mathbb{E}(Y|Z=0) = 0 \cdot f_{Y|Z}(0,0) + 1 \cdot f_{Y|Z}(1,0) = f_{Y|Z}(1,0) = 1$$

and

$$\mathbb{E}(Y|Z=1) = f_{Y|Z}(1,1) = \frac{b-a}{1-a}$$

so that

$$\mathbb{E}(Y|Z) = \begin{cases} 1 & Z=0 \\ \frac{b-a}{1-a} & Z=1. \end{cases}$$

(23) <u>Poisson(λ)</u>

$$\psi_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{e^{(t+\log\lambda)x}}{x!} = e^{-\lambda} e^{e^{t+\log\lambda}} = e^{\lambda(e^t-1)}$$

<u>$N(\mu,\sigma)$</u>

$$\psi_X(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{tx} e^{-(x-\mu)^2/2\sigma^2} dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{(-x^2+2\sigma^2 tx+2\mu x-\mu^2)/2\sigma^2} dx.$$

Completing the square yields

$$\psi_X(t) = e^{t\mu+\sigma^2t^2/2} \int \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\sigma^2 t-\mu)^2/2\sigma^2} dx = e^{t\mu+\sigma^2t^2/2}.$$

$\underline{\Gamma(\alpha)}$

$$\psi_X(t) = \int_0^\infty e^{tx} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx = \frac{1}{(1-t\beta)^\alpha} \int_0^\infty \frac{1}{\left(\frac{\beta}{1-t\beta}\right)^\alpha \Gamma(\alpha)} e^{-x(1-t\beta)/\beta} dx$$

and this is just equal to $\frac{1}{(1-t\beta)^\alpha}$ (assuming that $t < 1/\beta$).