solutions_for_hw_01

August 26, 2019

1 Problem 1

The number of ways to partition a set S of n elements into k subsets, $S_1, S_2, \ldots, S_k$, which are mutually exclusive and collectively exhaustive is given by the *Multinomial coefficient*

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!\, n_2!\, \ldots\, n_k!}, \tag{1.1}$$

where n_i is the number of elements of subset i and $n_1 + n_2 + \ldots + n_k = n$.

A proof of Eq. (1.1) is as follows.

There are $\binom{n}{n_1}$ ways of choosing subset S_1 from S. For each of the ways of choosing S_1, there are $\binom{n-n_1}{n_2}$ ways of choosing subset S_2 from the remaining elemnents in $S - S_1$, which denotes the set difference. We can reason in the same way until the last set S_k which has only 1 way to choose, given that we have chosen $S_1, S_2, \ldots, S_{k-1}$. So we have

$$\begin{aligned}\binom{n}{n_1, n_2, \ldots, n_k} &= \binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \ldots \times \binom{n-n_1-\ldots-n_{k-2}}{n_{k-1}} \times 1 \\ &= \frac{n!}{n_1!\,(n-n_1)!} \times \frac{(n-n_1)!}{n_2!\,(n-n_1-n_2)!} \times \ldots \times \frac{(n-n_1-\ldots-n_{k-2})!}{n_{k-1}!\,(n-n_1-\ldots-n_{k-2}-n_{k-1})!} \\ &= \frac{n!}{n_1!} \times \frac{1}{n_2!} \times \ldots \times \frac{1}{n_{k-1}!\,n_k!} \\ &= \frac{n!}{n_1!\,n_2!\,\ldots\,n_k!}. \end{aligned} \tag{1.2}$$

Q.E.D.

So the number of ways of dividing 15 students into 3 groups of sizes 4, 5 and 6, respectively, is:

$$\binom{15}{4, 5, 6} = \frac{15!}{4!\,5!\,6!} = 630, 630. \tag{1.3}$$

2 Problem 2

We can represent the box as a vector

$$b = \begin{bmatrix} R \\ R \\ G \\ B \end{bmatrix}, \tag{2.1}$$

where $R \equiv$ red marble, $G \equiv$ green marble and $B \equiv$ blue marble.

- If we draw the marbles *with replacement*, the first and second draws are from the same box, b. Therefore, the sample space in this case is the all possible pairs formed by the *outer product* of b with itself

$$\Omega_1 = bb^T = \begin{bmatrix} R \\ R \\ G \\ B \end{bmatrix} \begin{bmatrix} R & R & G & B \end{bmatrix} = \begin{bmatrix} RR & RR & RG & RB \\ RR & RR & RG & RB \\ GR & GR & GG & GB \\ BR & BR & BG & BB \end{bmatrix}. \tag{2.2}$$

- If we draw the marbles *without replacement*, the marble drawn the first time will not be available for the second draw. So we need to exlude pairs in which the first and the second marble are identical. The sample space in this case is Ω_1 minus the diagonal.

$$\Omega_2 = \Omega_1 - \operatorname{diag}\left(\begin{bmatrix} RR & RR & GG & BB \end{bmatrix}\right) = \begin{bmatrix} * & RR & RG & RB \\ RR & * & RG & RB \\ GR & GR & * & GB \\ BR & BR & BG & * \end{bmatrix}. \tag{2.3}$$

3 Problem 3

Let Ω be the sample space of rolling two dice:

$$\Omega = \begin{bmatrix} 1\,1 & 1\,2 & 1\,3 & 1\,4 & 1\,5 & 1\,6 \\ 2\,1 & 2\,2 & 2\,3 & 2\,4 & 2\,5 & 2\,6 \\ 3\,1 & 3\,2 & 3\,3 & 3\,4 & 3\,5 & 3\,6 \\ 4\,1 & 4\,2 & 4\,3 & 4\,4 & 4\,5 & 4\,6 \\ 5\,1 & 5\,2 & 5\,3 & 5\,4 & 5\,5 & 5\,6 \\ 6\,1 & 6\,2 & 6\,3 & 6\,4 & 6\,5 & 6\,6 \end{bmatrix}. \tag{3.1}$$

Let A be the event that "*at least one die lands on 5*". A contains elements in the 5th row and elements in the 5th column of Ω. So

$$P(A) = \frac{11}{36}. \tag{3.2}$$

Let S_i is the event that "*sum of two dice is i*", where $i = 3, 4, \ldots, 12$. We have

$$P(A \mid S_i) = \frac{P(A \cap S_i)}{P(S_i)}. \tag{3.3}$$

From the matrix Ω we see that A contains pairs of dice having sum in the range $[6, 11]$. So we have $P\left(A \cap S_j\right) = 0$ for $j \in \{3, 4, 5, 12\}$ which means $P(A \mid S_j) = 0$ for $j \in \{3, 4, 5, 12\}$.

- $S_6 = \{1\,5,\ 2\,4,\ 3\,3,\ 4\,2,\ 5\,1\}$, $A \cap S_6 = \{1\,5,\ 5\,1\} \Rightarrow$

$$P(A \mid S_6) = \frac{P(A \cap S_6)}{P(S_6)} = \frac{2/36}{5/36} = \frac{2}{5}. \tag{3.4}$$

- $S_7 = \{1\,6,\ 2\,5,\ 3\,4,\ 4\,3,\ 5\,2,\ 6\,1\}$, $A \cap S_7 = \{2\,5,\ 5\,2\} \Rightarrow$

$$P(A \mid S_7) = \frac{P(A \cap S_7)}{P(S_7)} = \frac{2/36}{6/36} = \frac{1}{3}. \tag{3.5}$$

- $S_8 = \{2\,6,\ 3\,5,\ 4\,4,\ 5\,3,\ 6\,2\}$, $A \cap S_8 = \{3\,5,\ 5\,3\} \Rightarrow$

$$P(A \mid S_8) = \frac{P\left(A \cap S_8\right)}{P(S_8)} = \frac{2/36}{5/36} = \frac{2}{5}. \tag{3.6}$$

- $S_9 = \{3\,6,\ 4\,5,\ 5\,4,\ 6\,3\}$, $A \cap S_9 = \{4\,5,\ 5\,4\} \Rightarrow$

$$P(A \mid S_9) = \frac{P\left(A \cap S_9\right)}{P(S_9)} = \frac{2/36}{4/36} = \frac{1}{2}. \tag{3.7}$$

- $S_{10} = \{4\,6,\ 5\,5,\ 6\,4\}$, $A \cap S_{10} = \{5\,5\} \Rightarrow$

$$P(A \mid S_{10}) = \frac{P\left(A \cap S_{10}\right)}{P(S_{10})} = \frac{1/36}{3/36} = \frac{1}{3}. \tag{3.8}$$

- $S_{11} = \{5\,6,\ 6\,5\}$, $A \cap S_{11} = \{5\,6,\ 6\,5\} \Rightarrow$

$$P(A \mid S_{11}) = \frac{P\left(A \cap S_{11}\right)}{P(S_{11})} = \frac{2/36}{2/36} = 1. \tag{3.9}$$

4 Problem 4

The PMF of a Poisson random variable X is given by

$$P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \tag{4.1}$$

where λ is the expected number of events occuring in an interval, and i is a non-negative interger.

Consider the ratio

$$\frac{P(X = i)}{P(X = i-1)} = \frac{\lambda^i/i!}{\lambda^{i-1}/(i-1)!} = \frac{\lambda}{i}. \tag{4.2}$$

- For all i such that $i < \lambda$, we have $\frac{\lambda}{i} > 1$ or $P(X = i) > P(X = i-1)$. This means $P(X = i)$ monotonically increases when $i < \lambda$.
- For all i such that $i > \lambda$, we have $\frac{\lambda}{i} < 1$ or $P(X = i) < P(X = i-1)$. This means $P(X = i)$ monotonically decreases when $i > \lambda$.

Now we need to prove that the maximun stays on the left side of λ (if λ is not an inetger). Let $i_L = \underset{i<\lambda}{\text{argmax}} P(X = i)$ and $i_R = \underset{i>\lambda}{\text{argmax}} P(X = i)$. Then we need to prove that $P(X = i_L) > P(X = i_R)$.

Since $P(X = i)$ monotonically increases on the left and monotonically decreases on the right of λ, we have $i_R = i_L + 1$ and $i_L < \lambda < i_R$. This mean that there exists $\rho \in (0, 1)$ such that $i_L = \lambda - \rho$ and $i_R = \lambda + (1 - \rho)$. Now let's check the ratio

$$\frac{P(X = i_L)}{P(X = i_R)} = \frac{P(X = \lambda - \rho)}{P(X = \lambda - \rho + 1)} = \frac{\lambda - \rho + 1}{\lambda} \tag{4.3}$$

Since $\rho < 1 \Rightarrow \frac{\lambda - \rho + 1}{\lambda} > 1$ or $P(X = i_L) > P(X = i_R)$.

If λ happens to be an integer, then $i_L = i_R = \lambda$.

Q.E.D.

5 Prolem 5

The standard deviation of a random variable X is given by

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)} \equiv \sigma. \tag{5.1}$$

Let's do a linear transformation of X:

$$Y = aX + b, \tag{5.2}$$

and ask what is the standard deviation of Y?

$$\mathrm{Var}[Y] = \mathrm{E}[Y^2] - \mathrm{E}[Y]^2, \tag{5.3}$$

where $\mathrm{E}[\cdot]$ denotes the expectation. Note that the expectation has a linearity property as follows

$$\mathrm{E}\left[(aX + b)\right] = a\mathrm{E}[X] + b. \tag{5.4}$$

We have

$$\begin{aligned}
\mathrm{Var}[Y] &= \mathrm{E}\left[(aX+b)^2\right] - \left(\mathrm{E}\left[(aX+b)\right]\right)^2 \\
&= \mathrm{E}[a^2X^2 + 2abX + b^2] - \left(a\mathrm{E}[X] + b\right)^2 \\
&= a^2\mathrm{E}[X^2] + 2ab\mathrm{E}[X] + b^2 - a^2\mathrm{E}[X]^2 - 2ab\mathrm{E}[X] - b^2 \\
&= a^2\left(\mathrm{E}[X^2] - \mathrm{E}[X]^2\right) \\
&= a^2\mathrm{Var}[X] = a^2\sigma^2.
\end{aligned} \tag{5.5}$$

So the standard deviation of Y is

$$\mathrm{SD}[Y] = \sqrt{\mathrm{Var}(Y)} = a\sigma. \tag{5.6}$$

6 Problem 6

6.1 (a)

X and Y are independent becuase we can factorize the PDF $f(x,y)$ into two functions, each only depends on one variable:

$$f(x,y) = h(x)g(y), \tag{6.1}$$

where we can choose $h(x) \equiv x(1-x)$ and $g(y) \equiv 12y$.

6.2 (b)

The marginal density of X is

$$f_X(x) = \int_0^1 f(x,y)dy = 12x(1-x)\int_0^1 ydy = 6x(1-x). \tag{6.2}$$

The expectation of X is

$$\mathrm{E}[X] = \int_0^1 x f_X(x) dx = 6 \int_0^1 x^2(1-x) dx = 6 \int_0^1 (x^2 - x^3) dx = \frac{1}{2}. \tag{6.3}$$

6.3 (c)

The marginal density of Y is

$$f_Y(y) = \int_0^1 f(x,y) dx = 12y \int_0^1 (x - x^2) dx = 2y. \tag{6.4}$$

The expectation of Y is

$$\mathrm{E}[Y] = \int_0^1 y f_Y(y) dy = 2 \int_0^1 y^2 dy = \frac{2}{3}. \tag{6.5}$$

6.4 (d)

$$\mathrm{E}[X^2] = \int_0^1 x^2 f_X(x) dx = \int_0^1 6x^3(1-x) dx = 6 \int_0^1 (x^3 - x^4) dx = \frac{3}{10}. \tag{6.6}$$

The variance of X is

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{3}{10} - \left(\frac{1}{2}\right)^2 = \frac{1}{20}. \tag{6.7}$$

6.5 (e)

$$\mathrm{E}[Y^2] = \int_0^1 y^2 f_Y(y) dy = 2 \int_0^1 y^3 dy = \frac{1}{2}. \tag{6.8}$$

The variance of Y is

$$\mathrm{Var}(Y) = \mathrm{E}[Y^2] - \mathrm{E}[Y]^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}. \tag{6.9}$$

7 Problem 7

Since we do not like to take risk, we should choose the scheme that has the lowest variance. So we will calculate variance of X and Y to see which random variable has a lower variance.

Let N be the number of dots on a die. Assuming that the die is fair, the expectation of N is

$$\mathrm{E}[N] = \sum_{i=1}^{6} iP(X = i) = \frac{1}{6} \sum_{i=1}^{6} i = 3.5. \tag{7.1}$$

To calculate the variance, we need to calculate $\mathrm{E}[N^2]$

$$\mathrm{E}[N^2] = \frac{1}{6} \sum_{i=1}^{6} i^2 = 15.1666. \tag{7.2}$$

$$\mathrm{Var}[N] = \mathrm{E}[N^2] - \mathrm{E}[N]^2 = 15.1666 - (3.5)^2 = 2.9166. \tag{7.3}$$

7.1 Scheme (a)

$$X = 1000 \times N. \tag{7.4}$$

$$\mathrm{E}[X] = 1000 \times \mathrm{E}[N] = 1000 \times 3.5 = 3500. \tag{7.5}$$

From Eq. (5.6) above, we have

$$\mathrm{Var}[X] = 1000^2 \times \mathrm{Var}[N] = 2.916 \times 10^6. \tag{7.6}$$

7.2 Scheme (b)

$$Y = \sum_{i=1}^{1000} N_i, \tag{7.7}$$

where $\{N_i\}$ are independent and identically distributed (iid) random variables. Each N_i follows a *multinoulli distribution* (*categorical distribution*) with six classes, each having the same probability of occuring ($\frac{1}{6}$). Due to the linearity property of the expectation $\mathrm{E}[\cdot]$, we have

$$\mathrm{E}[Y] = \mathrm{E}\left[\sum_{i=1}^{1000} N_i\right] = \sum_{i=1}^{1000} \mathrm{E}[N_i] = 1000 \times \mathrm{E}[N] = 1000 \times 3.5 = 3500. \tag{7.8}$$

So both schemes give the same expectation.

To calculate $\mathrm{Var}[Y]$, we will make use of the following equality which holds true for sum of independent random variables:

$$\mathrm{Var}\left[\sum_{i=1}^{N} a_i U_i\right] = \sum_{i=1}^{N} a_i^2 \mathrm{Var}[U_i]. \tag{7.9}$$

Let's prove Eq. (7.9). Let $\{U_i\}$ be a set on *independent random variables* with expectation $\mu_i \equiv \mathrm{E}[U_i]$ and variance $\sigma_i^2 \equiv \mathrm{Var}[U_i]$. We define a random variable Z which is a linear combination of $\{U_i\}$:

$$Z = \sum_{i=1}^{N} a_i U_i. \tag{7.10}$$

The expectation of Z is

$$\mu_Z \equiv \mathrm{E}[Z] = \sum_{i=1}^{N} a_i \mu_i. \tag{7.11}$$

The variance of Z is

$$
\begin{aligned}
\mathrm{Var}[Z] &= \mathrm{E}\left[(Z-\mu_Z)^2\right] \\
&= \mathrm{E}\left[\left(\sum_{i=1}^{N} a_i U_i - \sum_{i=1}^{N} a_i \mu_i\right)^2\right] \\
&= \mathrm{E}\left[\left(\sum_{i=1}^{N} a_i (U_i - \mu_i)\right)^2\right] \\
&= \mathrm{E}\left[\left(\sum_{i=1}^{N} a_i (U_i - \mu_i)\right)\left(\sum_{j=1}^{N} a_j (U_j - \mu_j)\right)\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j (U_i - \mu_i)(U_j - \mu_j)\right] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j \mathrm{E}\left[(U_i - \mu_i)(U_j - \mu_j)\right]. \qquad (7.12)
\end{aligned}
$$

For $i = j$, $\mathrm{E}\left[(U_i - \mu_i)(U_i - \mu_i)\right] = \sigma_i^2$ which is the variance of U_i. For $i \neq j$, $\mathrm{E}\left[(U_i - \mu_i)(U_j - \mu_j)\right] = \sigma_{i,j}^2$ which is the *covariance* of U_i and U_j. However, the covariance is zero, $\sigma_{i,j}^2 = 0$, becuase U_i and U_j are independent by definition. So the double sum in Eq.(7.12) becomes

$$
\mathrm{Var}[Z] = \mathrm{Var}\left[\sum_{i=1}^{N} a_i U_i\right] = \sum_{i=1}^{N} a_i^2 \mathrm{Var}[U_i] \qquad (7.13)
$$

Q.E.D.

Now using Eq. (7.13) to calculate $\mathrm{Var}[Y]$ we have

$$
\mathrm{Var}[Y] = \sum_{i=1}^{1000} \mathrm{Var}[N_i] = 1000 \times \mathrm{Var}[N] = 2.916 \times 10^3. \qquad (7.14)
$$

From Eqs (7.6) and (7.10) we see that $\mathrm{Var}[Y] < \mathrm{Var}[X]$. So we will choose **scheme (b)**.

7.3 Simulation approach to estimating the variance

We can also estimate $\mathrm{Var}[X]$ and $\mathrm{Var}[Y]$ using *simulation method* as follows.

```
In [1]: import numpy as np
```

```
In [5]: np.random.seed(1234)
        nsamples = 5000
        die = list(range(1, 7))
        print("The die:", die)

        x_samples = 1000 * np.random.choice(die, size=nsamples, replace=True)
        print("Variance of X (scheme a): %0.4e" % np.var(x_samples))

        y_samples = [sum(np.random.choice(die, size=1000, replace=True))
```

```python
                        for _ in range(nsamples)]
        print("Variance of Y (scheme b): %0.4e" % np.var(y_samples))
```

```
The die: [1, 2, 3, 4, 5, 6]
Variance of X (scheme a): 2.9235e+06
Variance of Y (scheme b): 2.8819e+03
```