Tutorial 9

November 26, 2020

Question 1

You toss a fair coin repeatedly.

(i) What is the expected number of tosses until the pattern HT appears for the first time?

Let W_{HT} be the number of tosses until the pattern HT appears for the first time.

$$\underbrace{TTTH}_{W_1}\ \underbrace{HH\boldsymbol{HT}}_{W_2}\ HHT\ldots$$

From the example above, we can see that W_{HT} can be thought of as the waiting time for the first heads, call this W_1, plus the additional waiting time for the first tails after the heads, call this W_2.

Then W_1 and W_2 are i.i.d. Geometric($p = \frac{1}{2}$), where W_1 and W_2 are counting the total number of trials. Then:

$$\mathbf{E}(W_1) = \mathbf{E}(W_2) = \frac{1}{p} = 2 \text{ trials}$$

As a result,

$$\mathbf{E}(W_{HT}) = \mathbf{E}(W_1 + W_2) = 2 + 2 = 4 \text{ trials}$$

(ii) What is the expected number of tosses until the pattern HH appears for the first time?

In observing the waiting time until the sequence HH appears for the first time, we cannot apply the same logic as the previous part.

$$T\underset{1}{H}TTTT\underset{1}{H}\underset{2}{H}HHT\ldots$$

From the example above, if the first heads is immediately followed by tails, we must start over and wait for the first heads to appear again. This does help us as the fact that the system resets suggests that we should consider a *first-step analysis*, i.e. condition on the first step of the experiment. Let us condition on the outcome of the first toss (we are actually using the law of total expectation!):

$$\begin{aligned}\mathbf{E}(W_{HH}) &= \mathbf{E}(W_{HH} \mid \text{first toss } H)\,\mathbf{P}(\text{first toss } H) + \mathbf{E}(W_{HH} \mid \text{first toss } T)\,\mathbf{P}(\text{first toss } T) \\ &= \mathbf{E}(W_{HH} \mid \text{first toss } H)\,\frac{1}{2} + \mathbf{E}(W_{HH} \mid \text{first toss } T)\,\frac{1}{2} \qquad (*)\end{aligned}$$

The second conditional expectation, $\mathbf{E}(W_{HH} \mid \text{first toss } T) = 1 + \mathbf{E}(W_{HH})$, due to lack of memory: since W_{HH} is counting the number of trials required to obtain the pattern HH, a first toss resulting in T does not help us in this pursuit and can be thought of as adding an additional trial to our expected waiting time.

For the first term, we compute $\mathbf{E}(W_{HH} \mid \text{first toss } H)$ by further conditioning on the outcome of the second toss. If the second toss is heads, we have obtained HH in two tosses. If the second toss is tails, we have wasted two tosses and have to start over. This gives:

$$\begin{aligned}\mathbf{E}(W_{HH} \mid \text{first toss } H) &= \mathbf{E}(W_{HH} \mid HH)\,\mathbf{P}(\text{second toss } H) + \mathbf{E}(W_{HH} \mid HT)\,\mathbf{P}(\text{second toss } T) \\ &= 2 \cdot \frac{1}{2} + (2 + \mathbf{E}(W_{HH})) \cdot \frac{1}{2} \\ &= 2 + \mathbf{E}(W_{HH}) \cdot \frac{1}{2}\end{aligned}$$

Revisiting $(*)$ and combining our results, gives:

$$\mathbf{E}(W_{HH}) = \left(2 + \mathbf{E}(W_{HH}) \cdot \frac{1}{2}\right) \cdot \frac{1}{2} + (1 + \mathbf{E}(W_{HH})) \cdot \frac{1}{2}$$

Solving for $\mathbf{E}(W_{HH})$ gives $\mathbf{E}(W_{HH}) = 6$.

Question 2

An immortal drunk man wanders around randomly on the integers. He starts at the origin, and at each step he moves 1 unit to the right or 1 unit to the left, with equal probabilities, independently of all his previous steps. Let b be a googolplex (that is, 10^g where $g = 10^{100}$ is a googol).

(a) Find a simple expression for the probability that the immortal drunk visits b before returning to the origin for the first time.

Let B be the event that the man visits b before returning to the origin for the first time. Let L be the event that his first move is to the left.

$\mathbf{P}(B \mid L) = 0$ for if his first move is to the left, he will end up at -1, and going from -1 to b must pass through the origin.

$\mathbf{P}(B \mid L^c)$ can be found using the result of the *Gambler's Ruin* problem (see page 64 of *Introduction to Probability* by Blitzstein and Hwang) where player A starts with i dollars, player B starts with $b-1$ dollars, and at each round each player has an equal probability of winning. Applying these results with the law of total probability, we obtain:

$$\mathbf{P}(B) = \mathbf{P}(B \mid L)\,\mathbf{P}(L) + \mathbf{P}(B \mid L^c)\,\mathbf{P}(L^c) = \frac{1}{b} \cdot \frac{1}{2} = \frac{1}{2b}$$

Since b in this case is extremely large, this probability is essentially zero.

(b) Find the expected number of times that the immortal drunk visits b before returning to the origin for the first time.

Let N be the number of visits to b before returning to the origin for the first time and let $p = \frac{1}{2b}$ from the result of the previous part. Applying the law of total expectation:

$$\mathbf{E}(N) = \mathbf{E}(N \mid N = 0)\,\mathbf{P}(N = 0) + \mathbf{E}(N \mid N \geq 1)\,\mathbf{P}(N \geq 1) = p \cdot \mathbf{E}(N \mid N \geq 1)$$

Note that $\mathbf{P}(N \geq 1) = \mathbf{P}(B) = p = \frac{1}{2b}$ since B was defined as the event that b is visited before returning to the origin and as such, the man can visit b numerous times, but at least one time, before returning to the origin. The number of times that he visits b before returning to the origin is captured by the probability of $N \geq 1$.

The conditional distribution of N given $N \geq 1$ is Geometric(p):

- given that the man reaches b, by symmetry, there is probability p of returning to the origin before visiting b again (denote this as "success") and probability $1 - p$ of returning to b again before returning to the origin (denote this as "failure")
- the trials are independent since the situation is the same each time he is at b, independent of the past history

Then $\mathbf{E}\left(N \mid N \geq 1\right) = \frac{1}{p}$ and

$$\mathbf{E}\left(N\right) = p \cdot \mathbf{E}\left(N \mid N \geq 1\right) = p \cdot \frac{1}{p} = 1$$

Interestingly, this result does not depend on the value of b and the value of p was not required to be known.

Question 3

Let $\boldsymbol{X} = (X_1, X_2, X_3)$ have a multivariate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Find $\mathbf{P}\left(X_1 > X_2 + X_3 + 2\right)$.

To find this probability, we can reduce our problem to a univariate normal probability.

$$\mathbf{P}\left(X_1 > X_2 + X_3 + 2\right) = \mathbf{P}\left(X_1 - X_2 - X_3 > 2\right)$$

Letting $U := X_1 - X_2 - X_3$, the above probability becomes:

$$\mathbf{P}\left(U > 2\right)$$

where U still follows a normal distribution.

$$\mathbf{E}\left(U\right) = \mathbf{E}\left(X_1 - X_2 - X_3\right) = 0 - 0 - 0 = 0$$

$$\begin{aligned}
\mathbf{Var}\left(U\right) &= \mathbf{Cov}\left(U,\, U\right) \\
&= \mathbf{Cov}\left(X_1 - X_2 - X_3,\, X_1 - X_2 - X_3\right) \\
&= \mathbf{Var}\left(X_1\right) + \mathbf{Var}\left(X_2\right) + \mathbf{Var}\left(X_3\right) \\
&\quad - 2\mathbf{Cov}\left(X_1,\, X_2\right) - 2\mathbf{Cov}\left(X_1,\, X_3\right) - 2\mathbf{Cov}\left(X_2,\, X_3\right) \\
&= 1 + 2 + 2 - 0 - 0 + 2 \\
&= 7
\end{aligned}$$

Therefore $U \sim N(0, 7)$.

$$\mathbf{P}\left(U > 2\right) \;=\; \mathbf{P}\left(Z > \frac{2-0}{\sqrt{7}}\right) \;=\; \mathbf{P}\left(Z > \frac{2}{\sqrt{7}}\right) \;=\; 0.2248$$

Question 4

Let X, Y, Z be i.i.d. $N(0,1)$. Find the joint MGF of

$$(X+2Y,\, 3X+4Z,\, 5Y+6Z)$$

Method 1: Using matrices

$$\mathbf{w} \;:=\; \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \;\sim\; \mathcal{N}_3\left(\mathbf{0},\, \mathbb{I}_3\right) \quad \text{(given in the question)}$$

$$\mathbf{v} \;=\; \begin{bmatrix} X+2Y \\ 3X+4Z \\ 5Y+6Z \end{bmatrix} \;=\; \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 4 \\ 0 & 5 & 6 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \;:=\; \mathbf{A}\mathbf{w}$$

By the properties of the multivariate normal distribution, if

$$\mathbf{w} \;\sim\; \mathcal{N}_p(\boldsymbol{\mu},\, \boldsymbol{\Sigma}) \quad \text{then} \quad \mathbf{A}\mathbf{w} \;\sim\; \mathcal{N}_q(\mathbf{A}\boldsymbol{\mu},\, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$$

where $\mathbf{A}$ is a $q \times p$ constant matrix of rank q. Applying this to our case,

$$\mathbf{w} \;\sim\; \mathcal{N}_3(\mathbf{0},\, \mathbb{I}_3) \quad \text{then} \quad \mathbf{v} = \mathbf{A}\mathbf{w} \;\sim\; \mathcal{N}_3(\mathbf{A}\cdot\mathbf{0},\, \mathbf{A}\mathbb{I}_3\mathbf{A}')$$

Clearly, $\mathbf{E}\left(\mathbf{v}\right) \;=\; \mathbf{0}$.

$$\begin{aligned}
\mathbf{Var}\left(\mathbf{v}\right) &= \mathbf{A}\mathbb{I}_3\mathbf{A}' \\
&= \mathbf{A}\mathbf{A}' \\
&= \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 4 \\ 0 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 2 & 0 & 5 \\ 0 & 4 & 6 \end{bmatrix} \\
&= \begin{bmatrix} 5 & 3 & 10 \\ 3 & 25 & 24 \\ 10 & 24 & 61 \end{bmatrix}
\end{aligned}$$

Then the joint MGF is found using:

$$M(\mathbf{t}) \;=\; \mathbf{E}\left(e^{\mathbf{t}'\mathbf{v}}\right) \;=\; \exp\left(\mathbf{t}'\boldsymbol{\mu}_{\mathbf{v}} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{t}\right), \quad \mathbf{t} = (t_1, t_2, t_3) \in \mathbb{R}^3$$

as $\mathbf{v}$ is multivariate normal. The first term in the exponent will be zero since all components of $\boldsymbol{\mu}_{\mathbf{v}}$ are zero. The second term in the exponent works out to be:

$$\frac{1}{2}(5t_1^2 + 25t_2^2 + 61t_3^2 + 6t_1t_2 + 20t_1t_3 + 48t_2t_3)$$

Method 2: Direct method

$$
\begin{aligned}
M(\mathbf{t}) &= \mathbf{E}\left(e^{\mathbf{t}'\mathbf{v}}\right) \\
&= \mathbf{E}\left(\exp\left(t_1(X+2Y) \;+\; t_2(3X+4Z) \;+\; t_3(5Y+6Z)\right)\right) \\
&= \mathbf{E}\left(\exp\left((t_1+3t_2)X \;+\; (2t_1+5t_3)Y \;+\; (4t_2+6t_3)Z\right)\right)
\end{aligned}
$$

Letting $R \;=\; (t_1+3t_2)X \;+\; (2t_1+5t_3)Y \;+\; (4t_2+6t_3)Z$,

$$
M(\mathbf{t}) = \exp\left(\mathbf{E}\left(R\right) \;+\; \frac{1}{2}\mathbf{Var}\left(R\right)\right), \quad \mathbf{t} = (t_1, t_2, t_3) \in \mathbb{R}^3
$$

as R has a normal distribution. Clearly, $\mathbf{E}\left(R\right) \;=\; 0$.

$$
\begin{aligned}
\frac{1}{2}\mathbf{Var}\left(R\right) &= \frac{1}{2}\left((t_1+3t_2)^2 \;+\; (2t_1+5t_3)^2 \;+\; (4t_2+6t_3)^2\right) \\
&= \frac{1}{2}(5t_1^2 + 25t_2^2 + 61t_3^2 + 6t_1t_2 + 20t_1t_3 + 48t_2t_3)
\end{aligned}
$$