Probability and Statistics Homework 11

Hairui Yin

1. A trader invented a strategy where he can either win or lose a certain amount of money every day (say, k dollars) with equal probability, independently of all other days. (In fact, such a strategy is not very difficult to implement - buy a very liquid asset whose price moves without major jumps and liquidate the position as soon as you are making or losing k dollars). He then made the following statement: "If I use my strategy every day for the next 400 days, my chances of ending up within 10,000 dollars of the original account balance are equal to 1/3.

(a) What is k (approximately)?

(b) What are his chances (approximately) of making over 20,000 with this strategy in 400 days?

Answer:

(a) For each step, the trader has 50% winning k dollars and 50% winning -k dollars, which follows a Bernoulli distribution, with mean 0, variance k^2, and $\sigma = k$. Since each round is i.i.d., let X be the total money after the game, $E[X] = 0, \sigma_X = \sqrt{n}\sigma = 20k$.

$$
\begin{aligned}
P(-10000 \leq X \leq 10000) &= P(\frac{-10000 - 0}{20k} \leq \frac{X - \mu}{\sigma_X} \leq \frac{10000 - 0}{20k}) \\
&= P(\frac{-500}{k} \leq Z \leq \frac{500}{k})
\end{aligned}
$$

Consider the z-score of $\Phi(a) - \Phi(-a) = 2\Phi(a) - 1 = \frac{1}{3}$, we can solve that $a \approx 0.43$. Therefore,

$$
\begin{aligned}
\frac{500}{k} &\approx 0.43 \\
\Rightarrow k &\approx 1162.79
\end{aligned}
$$

(b) The probability of making over 20,000 is

$$
\begin{aligned}
P(X \geq 20000) &= P(Z \geq \frac{20000 - 0}{20k}) \\
&= P(Z \geq \frac{20000}{20 * 1162.79}) \\
&= P(Z \geq 0.86) \\
&= 1 - \Phi(0.86) \\
&= 1 - 0.8051 \\
&= 0.1949
\end{aligned}
$$

Therefore, the trader has 0.1949 chances to make over 20,000 with his strategy.

2. You are trying to use a machine that only works on some days. If on a given day, the machine is working it will break down the next day with probability $0 < b < 1$, and works on the next day with probability $1 - b$. If it is not working on a given day, it will work on the next day with probability $0 < r < 1$ and not work the next day with probability $1 - r$.

(a) In this problem we will formulate this process as a Markov chain. First, let $X^{(t)}$ be a variable that denotes the state of the machine at time t. Then, define a state space S that includes all the possible states that the machine can be in. Lastly, for all $A, B \in S$ find $P(X^{(t+1)} = A \mid X^{(n)} = B)$ (A and B can be the same state).

(b) Suppose that on day 1, the machine is working. What is the probability that it is working on day 3?

Answer:

(a) Let W represent the machine is working, NW represent the machine is not working. The state space is $S = \{W, NW\}$.
Given the probability information in the question, for adjacent states, the transition is defined as

	W	NW
W	$1-b$	b
NW	r	$1-r$

Thus the transition matrix is

$$Q = \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix}$$

Since $A, B \in S = \{W, NW\}$, we define a mapping function f, where $f(S) = 1$ if $S = W$, $f(S) = 2$ if $S = NW$.
Therefore, the probability is

$$P(X^{(t+1)} = A \mid X^{(n)} = B) = (f(B), f(A))^{th} \text{ entry of } Q^{t+1-n}$$

(b) To find this, we can calculate the two-step transition probability Q^2.

$$\begin{aligned} Q^2 &= QQ \\ &= \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix} \times \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix} \\ &= \begin{bmatrix} (1-b)^2 + br & b(1-b) + b(1-r) \\ r(1-b) + r^2 & br + (1-r)^2 \end{bmatrix} \end{aligned}$$

Given the result of question (a), the probability that day 1 machine working and day 3 machine working is $P = (1-b)^2 + br$.

3. Suppose that $X_1, ..., X_4$ are independent normal random variables with parameters (a, σ^2). Find the density of the random variable $Y = X_1 - 2X_2 + 3X_3 - 4X_4$.

Answer:

$X_1, ..., X_4$ are i.i.d. normal random variables. Since we know that Y also follows the normal distribution because its the linear combinition of X, we can find it mean and variance.

$$\begin{aligned} E[Y] &= E[X_1 - 2X_2 + 3X_3 - 4X_4] \\ &= E[X_1] - E[2X_2] + E[3X_3] - E[4X_4] \\ &= E[X_1] - 2E[X_2] + 3E[X_3] - 4E[X_4] \\ &= a - 2a + 3a - 4a \\ &= -2a \end{aligned}$$

$$\begin{aligned} Var(Y) &= Var(X_1 - 2X_2 + 3X_3 - 4X_4) \\ &= Var(X_1) + Var(2X_2) + Var(3X_3) + Var(4X_4) \\ &= Var(X_1) + 2^2 Var(X_2) + 3^2 Var(X_3) + 4^2 Var(X_4) \\ &= \sigma^2 + 4\sigma^2 + 9\sigma^2 + 16\sigma^2 \\ &= 30\sigma^2 \end{aligned}$$

Therefore, $Y \sim \mathcal{N}(-2a, 30\sigma^2)$. The density function is

$$f_Y(y) = \frac{1}{\sqrt{60\pi\sigma^2}} e^{\frac{-(y+2a)^2}{60\sigma^2}}$$