Probability theory 1

Class Test 2
B. Stat 1st year
Maximum marks: 20
Time: 1 hour

December 22, 2021, 11:30 A.M. - 12:30 P.M.

1. Let $(S_n : n \geq 0)$ be a simple symmetric random walk starting from 0. Define, for a fixed n,

$$M := \max\{S_0, \ldots, S_{2n}\}\,.$$

Show that

$$\mathrm{E}(M|S_{2n}=0) = \frac{1}{2}\left(\binom{2n}{n}^{-1} 4^n - 1\right)\,.$$

Soln.: Since M takes values $0, 1, 2, \ldots, n$ given that $S_{2n} = 0$,

$$\begin{aligned}
\mathrm{E}(M|S_{2n}=0) &= \sum_{k=1}^{n} P(M \geq k | S_{2n} = 0) \\
&= \binom{2n}{n}^{-1} 4^n \sum_{k=1}^{n} P\left(M \geq k, S_{2n} = 0\right) \\
\text{(Reflection principle)} &= \binom{2n}{n}^{-1} 4^n \sum_{k=1}^{n} P\left(S_{2n} = -2k\right) \\
&= \binom{2n}{n}^{-1} \sum_{k=1}^{n} \binom{2n}{n-k} \\
\text{(putting } j = n-k\text{)} &= \binom{2n}{n}^{-1} \sum_{j=0}^{n-1} \binom{2n}{j} \\
&= \frac{1}{2}\binom{2n}{n}^{-1} \sum_{j\in\{0,\ldots,2n\}\setminus\{n\}} \binom{2n}{j} \\
&= \frac{1}{2}\binom{2n}{n}^{-1} \left(4^n - \binom{2n}{n}\right) \\
&= \frac{1}{2}\left(\binom{2n}{n}^{-1} 4^n - 1\right)\,.
\end{aligned}$$

2. A random sample of size n is drawn from $\{1, \ldots, N\}$ with replacement. Calculate the variance of the number of distinct units in the sample.

Soln.: Define

$$X_i = \mathbf{1}\left(i \text{ is selected in the sample}\right), 1 \le i \le N\,.$$

If Z is the number of distinct units selected, then

$$Z = \sum_{i=1}^N X_i\,.$$

Thus,

$$\begin{aligned}\mathrm{Var}(Z) &= \sum_{i=1}^N \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le N} \mathrm{Cov}(X_i, X_j)\\ &= N\left(\frac{N-1}{N}\right)^n\left[1-\left(\frac{N-1}{N}\right)^n\right] + 2\sum_{1\le i<j\le N}\mathrm{Cov}(X_i, X_j)\,,\end{aligned}$$

because for all i, X_i follows Bernoulli with parameter $1-\left(\frac{N-1}{N}\right)^n$.

For $1 \le i < j \le N$,

$$\begin{aligned}&\mathrm{E}\left(X_iX_j\right)\\ &= P\left(\text{The sample contains both } i, j\right)\\ &= 1 - P\left(\text{The sample doesn't contain at least one of } i, j\right)\\ &= 1 - \Big[P\left(\text{The sample doesn't contain } i\right)\\ &\quad + P\left(\text{The sample doesn't contain } j\right)\\ &\quad - P\left(\text{The sample contains neither of } i, j\right)\Big]\\ &= 1 - \left[2\left(\frac{N-1}{N}\right)^n - \left(\frac{N-2}{N}\right)^n\right].\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) &= 1 - \left[2\left(\frac{N-1}{N}\right)^n - \left(\frac{N-2}{N}\right)^n\right] - \left[1-\left(\frac{N-1}{N}\right)^n\right]^2\\ &= \left(\frac{N-2}{N}\right)^n - \left(\frac{N-1}{N}\right)^{2n}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{Var}(Z) &= N\left(\frac{N-1}{N}\right)^n\left[1-\left(\frac{N-1}{N}\right)^n\right]\\ &\quad + N(N-1)\left[\left(\frac{N-2}{N}\right)^n - \left(\frac{N-1}{N}\right)^{2n}\right].\end{aligned}$$

3. A fair die is rolled till the first 6 appears. Calculate the expected number of times 1 is obtained.

Soln. 1.: For $i = 1, 2, \ldots, 5$, let X_i denote the number of times i is obtained. Clearly,

$$X_1 + \ldots + X_5 + 1 = Y \, .$$

Thus,

$$\sum_{i=1}^{5} \mathrm{E}(X_i) = \mathrm{E}(Y) - 1 = 5 \, ,$$

because Y follows Geometric(1/6) and hence $\mathrm{E}(Y) = 6$. Since $X_1, \ldots$ $\ldots, X_5$ have the same distribution, their expectations are the same as well, and therefore

$$\mathrm{E}(X_1) = 1 \, .$$

Soln. 2.: Define a roll of a die to be "valid" if either 1 or 6 turns up. In each valid roll of the die, 1 or 6 is obtained, each with probability 1/2. Let Z denote the number of valid rolls needed to get the first 6. Then Z follows Geometric(1/2). Clearly, the number of times 1 is obtained, in the original experiment, equals $Z - 1$. Thus,

$$\begin{aligned} \text{Expected number of times 1 is obtained} &= \mathrm{E}(Z - 1) \\ &= 2 - 1 \\ &= 1 \, . \end{aligned}$$