THE UNIVERSITY OF SYDNEY
SCHOOL OF MATHEMATICS AND STATISTICS

STAT4528 PROBABILITY AND MARTINGALE THEORY
SEMESTER 1 2023

LECTURERS: ANNA AKSAMIT AND BEN GOLDYS

DIAGNOSTIC QUIZ

Questions 1. Prove that the limit

$$\lim_{x\to 0} \sin \frac{1}{x}$$

does not exist.

Question 2. Let $E_1, \ldots, E_n$ be finite subsets of the set Ω. We denote by $|E|$ the number of elements in a finite set E. Prove the following

(1) $|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$

(2) $|E_1 \cup E_2 \cup E_3| = |E_1| + |E_2| + |E_3| - |E_1 \cap E_2| - |E_1 \cap E_3| - |E_2 \cap E_3| + |E_1 \cap E_2 \cap E_3|$.

(3) Using mathematical induction prove that the following holds true:

$$|E_1 \cup E_2 \cup \ldots \cup E_n| = \sum_{k=1}^{n} |E_k| - \sum_{i,j\in\{1,\ldots,n\}} |E_i \cap E_j| + \ldots + (-1)^{n+1} |E_1 \cap E_2 \cap \ldots \cap E_n|.$$

Question 3. Let X be a binomially distributed random variable with parameters $0 < p < 1$ and $n \geq 1$. This means that $Prob(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, \ldots, n$. Prove that

(1) $E(X) = np$,
(2) $Var(X) = np(1-p)$,
(3) Let $Y_1, \ldots, Y_n$ be independent Bernoulli random variables with parameter p, i.e., $Prob(Y_k = 1) = p$, and $Prob(Y_k = 0) = 1 - p$ for $k = 1, \ldots, n$. Show that the random variable $Z = Y_1 + \ldots + Y_n$ is binomially distributed with parameters p and n.
(4) Let $Y_1, \ldots$ be an infinite sequence of independent Bernoulli random variables with parameter p. Let

$$T = \min\{n \geq 1;\, Y_n = 1\}$$

if such an n exists and we put $T = \infty$ otherwise. Show that

$$Prob(T = \infty) = 0\,.$$

ANSWERS

Question 1. By definition $\lim_{x\to x_0} f(x)$ exists if there exists a such that for every sequence (x_n), $x_n \to x_0$, we have

$$\lim_{x_n\to x_0} f(x_n) = a\,.$$

Therefore, to show that the limit does not exist, it is enough to find two sequences (x_n) and (y_n), such that $x_n \to x_0$, $y_n \to x_0$, the limits $\lim_{n\to\infty} f(x_n)$ and $\lim_{n\to\infty} f(y_n)$ exist and

$$\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(x_n)\ .$$

To this end define

$$x_n = \frac{1}{\frac{\pi}{2}+2n\pi}, \quad y_n = \frac{1}{\frac{3\pi}{2}+2n\pi}.$$

Then $x_n \to 0$, $y_n \to 0$ and for every $n \geq 1$

$$-1 = \sin\frac{1}{y_n} \neq \sin\frac{1}{x_n} = 1\,.$$

Since both sequences are constant we obtain

$$-1 = \lim_{n\to\infty}\sin\frac{1}{y_n} \neq \lim_{n\to\infty}\sin\frac{1}{x_n} = 1\,.$$

Question 2.

(1) Assume first, that $E_1 \cap E_2 = \emptyset$. Then, the formula holds: $|E_1 \cup E_2| = |E_1| + |E_2|$ and a similar formula obviously holds for three disjoint sets:

$$|E_1 \cup E_2 \cup E_3| = |E_1| + |E_2| + |E_3|\,. \tag{1}$$

Since we have disjoint union

$$E_1 \cup E_2 = [(E_1 \setminus (E_1 \cap E_2))] \cup [(E_2 \setminus (E_1 \cap E_2))] \cup (E_1 \cap E_2)\,,$$

formula (1) gives

$$|E_1 \cup E_2| = |E_1| + |E_2| + |E_1 \cap E_2|\,.$$

Moreover,

$$E_1 = (E_1 \setminus (E_1 \cap E_2)) \cup (E_1 \cap E_2)\,,$$

and

$$E_2 = (E_2 \setminus (E_1 \cap E_2)) \cup (E_1 \cap E_2\,,$$

hence for $k = 1, 2$

$$|E_k| = |E_k \setminus (E_1 \cap E_2)| + |E_1 \cap E_2|\,,$$

and the claim follows immediately.Of course the proof follows easily from the geometric interpretation using Venn diagrams.

(2) Note that

$$E_1 \cup E_2 \cup E_3 = (E_1 \cup E_2) \cup E_3\,.$$

Therefore, using twice part 1 of the Question we obtain

$$\begin{aligned}|E_1 \cup E_2 \cup E_3| &= |E_1 \cup E_2| + |E_3| - |(E_1 \cup E_2) \cap E_3| \\ &= |E_1| + |E_2| + |E_3| - |E_1 \cap E_2| - |(E_1 \cup E_2) \cap E_3|\,.\end{aligned} \tag{2}$$

Using the fact that

$$(E_1 \cup E_2) \cap E_3 = (E_1 \cap E_3) \cup (E_2 \cap E_3)$$

and

$$(E_1 \cap E_3) \cap (E_2 \cap E_3) = E_1 \cap E_2 \cap E_3\,,$$

and invoking part 1 of the Question again we find that

$$|(E_1 \cup E_2) \cap E_3| = |E_1 \cup E_3| + |E_2 \cap E_3| - |E_1 \cap E_2 \cap E_3||$$

Inserting the last fromula into equation (2) we complete the proof.

(3) Use the same idea as in the proof of part 2, noting that for $n \geq 2$

$$E_1 \cup \cdots \cup E_n = (E_1 \cup \cdots E_{n-1}) \cup E_n .$$

Question 3.

(1) Note first that for $k \geq 1$

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} \tag{3}$$

Therefore

$$\begin{aligned} E(X) &= \sum_{k=0}^{n} k Prob(X = k) = \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} k\binom{n}{k} p^k (1-p)^{n-k} \\ &= np\sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1}(1-p)^{(n-1)-(k-1)} \\ &= np\sum_{k=0}^{n-1} \binom{n-1}{k} p^{k}(1-p)^{(n-1)-k} \\ &= np. \end{aligned}$$

The last inequality follows since $\binom{n-1}{k}p^k(1-p)^{(n-1)-k}$ are the binomial probabilities for a binomially distributed random variable with parameters p and $(n-1)$, hence

$$\sum_{k=0}^{n-1} \binom{n-1}{k} p^{k}(1-p)^{(n-1)-k} = 1 .$$

(2) Let us recall first that

$$Var(X) = E(X^2) - [E(X)]^2 ,$$

and since for $n = 1$ we have $X^2 = X$, the formula follows from part 1. We need to consider $n \geq 2$. Next

$$E(X^2) = E[X(X-1)] + E(X)$$

so that

$$Var(X) = E[X(X-1)] - [E(X)]^2 + E(X) \tag{4}$$

. It remains to compute $E[X(X-1)]$.We will show that

$$E[X(X-1)] = n(n-1)p^2 . \tag{5}$$

Using twice equation (3) we obtain for $n \geq k \geq 2$

$$\binom{n}{k} = \frac{n(n-1)}{k(k-1)}\binom{n-2}{k-2} ,$$

hence

$$\begin{aligned} E[X(X-1)] &= \sum_{k=2}^{n} k(k-1)P(X=k) \\ &= \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} \\ &= n(n-1)p^2\sum_{k=2}^{n}\binom{n-2}{k-2}p^{k-2}(1-p)^{(n-2)-(k-2)} \\ &= n(n-1)p^2\sum_{k=0}^{n-2}\binom{n-2}{k}p^{k}(1-p)^{(n-2)-k} \\ &= n(n-1)p^2\,. \end{aligned}$$

Combining (4), part 1 and (5) we obtain

$$Var(X) = n(n-1)p^2 - n^2p^2 + np = np(1-p)$$

as desired.

(3) Fix integers $1 \le i_1 < \cdots < i_k \le n$ and let $1 \le j_1 < \cdots < j_n - k \le n$ denote the remaing integers in the set $\{1,\ldots,n\}$. If $k=0$ or $k=n$ then one of these sets is empty. By independence

$$P(Y_{i_1} = \cdots = Y_{i_k} = 1, Y_{j_1} = \cdots = Y_{j_{n-k}} = 0) = p^k(1-p)^{n-k}\,.$$

Then number of choices of k indices $i_1,\ldots,i_k$ is equal to $\binom{n}{k}$ Then we have

$$\begin{aligned} P(Z=k) &= \sum_{\text{all choices of } \mathrm{i}_1\ldots,\mathrm{i}_\mathrm{k}} P(Y_{i_1} = \cdots = Y_{i_k} = 1, Y_{j_1}\cdots = Y_{j_{n-k}} = 0) \\ &= \sum_{\text{all choices of } \mathrm{i}_1\ldots,\mathrm{i}_\mathrm{k}} p^k(1-p)^{n-k} \\ &= \binom{n}{k}p^k(1-p)^{n-k} \end{aligned}$$

occurs if and only if k out of n random variables take value 1 and the remaining $(n-k)$ random variables take value 0. Eac

(4) It is enough to show that

$$P(T<\infty) = 1\,.$$

We have

$$P(T<\infty) = \sum_{k=1}^{\infty} P(T=k)\,.$$

Clearly, $P(T=1) = P(Y_1=1) = p$ and for $k \ge 2$ using independence we find that

$$P(T=k) = P(Y_1 = \cdots = Y_{k-1} = 0, Y_k = 1) = (1-p)^{k-1}p\,.$$

Therefore,

$$P(T<\infty) = \sum_{k=1}^{\infty}(1-p)^{k-1}p = 1\,.$$