Preliminary Assignment

CCMVI2085U Machine Learning for Predictive Analytics in Business

Bowei Chen

Copenhagen Business School ISUP 2020

I will review the business mathematics in lecture 02 and please take some time to complete this preliminary assignment before that. **It should be noted that this preliminary assignment is NOT included in your course assessment. It is mainly for you to recall the mathematical foundations as well as self-evaluate so that I can have a better understanding of your mathematical background for the course.** If you have any questions or need further advice, please get in touch with me at bc.acc@cbs.dk.

1. Roll a die and then flip that number of fair coins. What is the probability of "We get exactly 3 heads"?

   Solution

   Let A be the event that "We get exactly 3 heads". Let B_i="the die shows i", $\mathbb{P}(B_i) = 1/6$ for $i = 1, 2, \cdots, 6$ and

$$\mathbb{P}(A|B_1) = 0, \qquad \mathbb{P}(A|B_2) = 0, \qquad \mathbb{P}(A|B_3) = 2^{-3},$$

$$\mathbb{P}(A|B_4) = \binom{4}{3}2^{-4}, \qquad \mathbb{P}(A|B_5) = \binom{5}{3}2^{-5}, \qquad \mathbb{P}(A|B_6) = \binom{6}{3}2^{-6}.$$

   Then we have

$$\mathbb{P}(A) = \sum_{i=1}^{6} \mathbb{P}(A|B_i)\mathbb{P}(B_i) = \frac{1}{6}\left(\frac{1}{8} + \frac{4}{16} + \frac{10}{32} + \frac{20}{64}\right) = \frac{1}{6}.$$

2. If a random variable X follows a Poisson distribution $X \sim \text{Poi}(\lambda)$, please prove the following equations:

   (a) $\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1$;

   (b) $\mathbb{E}(X) = \lambda$;

   (c) $\text{Var}(X) = \lambda$.

   Solution

   (a)

$$\begin{aligned}\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} &= e^{-\lambda}\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \quad \text{(by Taylor expansion/theorem)} \\ &= e^{-\lambda}e^{\lambda} = 1.\end{aligned}$$

   (b)

$$\mathbb{E}(X) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}x = \lambda\sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = \lambda.$$

(c)

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\
&= \mathbb{E}(X^2) - \lambda^2 \\
&= \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} x^2 - \lambda^2 \\
&= \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} x(x-1) + \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} x - \lambda^2 \\
&= = \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} + \lambda - \lambda^2 \\
&= \lambda.
\end{aligned}
$$

3. If a random variable X follows a Binomial distribution $X \sim \mathrm{Bin}(n, p)$. Let $\lambda = np$. Suppose that n is very large and p is very small, please prove that $X \sim \mathrm{Poi}(\lambda)$.
Solution

$$
\begin{aligned}
\lim_{n\to\infty} f(x; n, p) &= \lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \lim_{n\to\infty} \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \lim_{n\to\infty} \left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-x+1}{n}\right) \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}.
\end{aligned}
$$

Since $\lim_{n\to\infty} \frac{n-x+1}{n} = 1$, $\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-x} = 1$, and $\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$, then

$$
\lim_{n\to\infty} f(x; n, p) = \frac{\lambda^x}{x!} e^{-\lambda} = f(x; \lambda).
$$

4. If a random variable X follows an exponential distribution, $X \sim \mathrm{Exp}(\lambda)$, please prove the following equations::
(a) $\int_0^\infty \lambda e^{-\lambda x} dx = 1$;
(b) $\mathbb{E}(X) = \frac{1}{\lambda}$;
(c) $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.
Solution
(a)

$$
\int_0^\infty \lambda e^{-\lambda x} dx = -\int_0^\infty de^{-\lambda x} = -(0-1) = 1.
$$

(b)

$$
\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty \lambda e^{-\lambda x} x dx \\
&= -\int_0^\infty x de^{-\lambda x} \\
&= -\left(x e^{-\lambda x}\Big|_0^\infty - \int_0^\infty e^{-\lambda x} dx \right) \\
&= \frac{1}{\lambda}.
\end{aligned}
$$

(c)

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \\
&= \mathbb{E}(X^2) - \frac{1}{\lambda^2} \\
&= \int_0^\infty \lambda e^{-\lambda x} x^2 dx - \frac{1}{\lambda^2} \\
&= -\int_0^\infty x^2 de^{-\lambda x} - \frac{1}{\lambda^2} \\
&= -\left(x^2 e^{-\lambda x} \Big|_0^\infty - \int_0^\infty e^{-\lambda x} dx^2 \right) - \frac{1}{\lambda^2} \\
&= \frac{2}{\lambda} \int_0^\infty \lambda x e^{-\lambda x} dx - \frac{1}{\lambda^2} \\
&= \frac{1}{\lambda^2}.
\end{aligned}
$$

5. Let $x_1, \cdots, x_n$ be a set of independently and identically distributed (i.i.d). random variables and each variable has a population mean μ and a finite variance σ^2. Please prove the following limiting behaviour:

$$
\lim_{n \to \infty} \sqrt{n} \left(\frac{\sum_{i=1}^n x_i}{n} - \mu \right) \sim N(0, \sigma^2).
$$

Solution
Let $Y = \sqrt{n}\big(\frac{\sum_{i=1}^n x_i}{n} - \mu\big)$ and let $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$, so $Y = \sigma Z$. According to the Central Limit Theorem, we can know $Z \sim N(0,1)$, then $Y \sim N(0, \sigma^2)$, completing the proof.

6. Prove that the sample variance $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{X})^2$ is an unbiased estimator, i.e., $\mathbb{E}(s^2) = \sigma^2$.

Solution
According to the definition of sample variance, we have

$$
\begin{aligned}
s^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2 \\
&= \frac{1}{n-1} \sum_{i=1}^n (x_i^2 - 2\bar{X}x_i + \bar{X}^2) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - 2n\bar{X}^2 + n\bar{X}^2 \right) \\
&= \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{X}^2 \right) \\
&= \frac{1}{n-1} \sum_{i=1}^n x_i^2 - \frac{1}{n(n-1)} \left(\sum_{i=1}^n x_i \right)^2
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathbb{E}(s^2) &= \frac{1}{n-1}\mathbb{E}\Big[\sum_{i=1}^n x_i^2\Big] - \frac{1}{n(n-1)}\mathbb{E}\Big[\Big(\sum_{i=1}^n x_i\Big)^2\Big] \\
&= \frac{1}{n-1}\sum_{i=1}^n \mathbb{E}[x_i^2] - \frac{1}{n(n-1)}\mathbb{E}\Big[\Big(\sum_{i=1}^n x_i\Big)^2\Big] \\
&= \frac{1}{n-1}\sum_{i=1}^n \Big(\mathrm{Var}(x_i) + \mathbb{E}^2[x_i]\Big) - \frac{1}{n(n-1)}\Big(\mathrm{Var}\Big[\sum_{i=1}^n x_i\Big] + \mathbb{E}^2\Big[\sum_{i=1}^n x_i\Big]\Big) \\
&= \frac{1}{n-1}\sum_{i=1}^n \Big(\sigma^2 + \mu^2\Big) - \frac{1}{n(n-1)}\Big(n\sigma^2 + n^2\mu^2\Big) \\
&= \sigma^2.
\end{aligned}
$$

7. From past experience, a professor knows that the test score of a student taking his final examination is a continuous random variable with mean 75.
   (a) Give an upper bound for the probability that a student's score will be at least as large as 85.
   (b) Suppose, in addition, that the professor knows that the variance of a student's test score is equal to 25. Give a lower bound to the probability that a student will score between 65 and 85?

   Solution

   (a) By Markov's inequality,

   $$\mathbb{P}(X \geq 85) \leq \mathbb{E}[X]/85 = 75/85 = 15/17.$$

   (b) By Chebyshev's inequality,

   $$\mathbb{P}(65 < X < 85) = 1 - \mathbb{P}(|X - 75| \geq 10) \geq 1 - 25/100 = 75/100 = 3/4,$$

   where we used that $\mathbb{P}(|X - 75| \geq 10) \leq \mathrm{Var}[X]/10^2 = 25/100 = 1/4$.
8. Coin number one comes up heads with probability 0.6 and coin number two with probability 0.5. A coin is continually flipped until it comes up tails, at which time that coin is put aside and we start flipping the other one.
   (a) What proportion of flips use coin 1?
   (b) If we start the process with coin 1, i.e., flip 1 uses coin 1, what is the probability that coin 2 is used on the fifth flip?

   Solution

   (a) Define state 1 to flip coin number one, and state 2 to flip coin number two. The Markov chain is $\{X_n, n \geq 1\}$ where X_n is the coin number of flip n. The reason why this is a Markov chain is that: (i) the number of states is finite, i.e., the random variable X_n is discrete, and (ii) the Markov property holds since the future evolution of the states (the coin number on each flip), is independent of the outcome of past flips if we condition on the current coin number (the current state). The transition probability matrix of the Markov chain is given by:

   $$\mathbf{P} = \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$$

   where we note that e.g., $\mathbf{P}_{1,1}$ is the probability that you get "head" next on the first coin, which occurs with probability 0.6. We find the stationary probabilities by solving:

   $$
   \begin{aligned}
   \pi_1 &= 0.6\pi_1 + 0.5\pi_2 \\
   \pi_2 &= 0.4\pi_1 + 0.5\pi_2 \\
   \pi_1 + \pi_2 &= 1,
   \end{aligned}
   $$

where we note that the second equation is redundant. By solving this system of equations, we get that $\pi_1 = \frac{5}{9}$ and $\pi_2 = \frac{4}{9}$. This means that we will use coin number one in $\frac{5}{9}$ of the flips.

(b) We have to find $\mathbb{P}(X_5 = 2 | X_1 = 1) = \mathbf{P}^4_{12}$. For this, we calculate:

$$\mathbf{P}^2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix}.$$

Also,

$$\mathbf{P}^4 = \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix} \cdot \begin{bmatrix} 0.56 & 0.44 \\ 0.55 & 0.45 \end{bmatrix} = \begin{bmatrix} 0.5556 & 0.4444 \\ 0.5556 & 0.4444 \end{bmatrix},$$

so that $\mathbf{P}^4_{12} \approx 0.44$.