THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH4240 - Stochastic Processes - 2020/21 Term 2

Homework 1
Due date: January 22, 2021

Please hand in your answers on Blackboard to all questions below.

Q1. A plane is missing, and it is presumed that it was equally likely to have gone down in any of 3 possible regions. Let $1 - \beta_i$, $i = 1, 2, 3$, denote the probability that the plane will be found upon a search of the i-th region when the plane is, in fact, in that region. (β_i: overlook probability). What is the conditional probability that the plane is in the i-th region given that a search of region 1 is unsuccessful?

Q2. Consider a random variable X taking the values

$$k_1, k_2, \cdots, k_n \in \mathbb{R}$$

with probability

$$p_1, p_2, \cdots, p_n \in [0, 1]$$

respectively, where $p_1 + p_2 + \cdots + p_n = 1$. Write down the formula for the expected value of $f(X)$ for a given function $f(\cdot)$.

Q3. Exercises of textbook (Chapter 1, starting from page 41): 4.

Q4. Compute the distribution of $X + Y$ in the following cases:

(a) X and Y are independent binomial random variables with parameters (n, p) and (m, p).

(b) X and Y are independent Poisson random variables with means respective λ_1 and λ_2.

(c) X and Y are independent normal random variables with respective parameters (μ_1, σ_1^2) and (μ_2, σ_2^2).

Q5. Read materials on *Law of Large Number* and *Central Limit Theorem* in the book "*A First Course in Probability*" by Ross (Chapter 8), and write down the statements of both theorems.

Solution:

Q1. Let E_i be the event that the plane is in the i-th region, and F be the event that a search of region 1 is unsuccessful. The required conidtional probability is $P(E_i|F)$ for $i = 1, 2$ and 3. By Bayes' formula,

$$P(E_i|F) = \frac{P(F|E_i)P(E_i)}{\sum_{k=1}^{3} P(F|E_k)P(E_k)}.$$

Since $P(E_k) = \frac{1}{3}$ for each k and

$$P(F|E_k) = \begin{cases} 1, & \text{if } k \neq 1, \\ \beta_1, & \text{if } k = 1. \end{cases}$$

We have

$$P(E_i|F) = \begin{cases} \dfrac{1}{2+\beta_1}, & \text{if } i \neq 1, \\ \dfrac{\beta_1}{2+\beta_1}, & \text{if } i = 1. \end{cases}$$

Q2. The random variable $f(X)$ takes the values

$$f(k_1), f(k_2), \ldots, f(k_n) \in \mathbb{R}$$

with probability $p_1, p_2, \ldots, p_n \in [0, 1]$ respectively. Therefore, the expected value $E[f(X)]$ of $f(X)$ is given by

$$E[f(X)] = \sum_{i=1}^{n} f(k_i)p_i.$$

Q3. (a)

$$\begin{aligned} P(C|\bigcup_i D_i) &= \frac{P(C \cap (\cup_i D_i))}{P(\cup_i D_i)} = \frac{P(\cup_i (C \cap D_i))}{P(\cup_i D_i)} \\ &= \frac{\sum_i P(C \cap D_i)}{\sum_i P(D_i)} = \frac{\sum_i P(C|D_i)P(D_i)}{\sum_i P(D_i)} \\ &= \frac{\sum_i p \cdot P(D_i)}{\sum_i P(D_i)} = p \end{aligned}$$

(b)

$$\begin{aligned} P(\bigcup_i C_i|D) &= \frac{P((\cup_i C_i) \cap D)}{P(D)} = \frac{P(\cup_i (C_i \cap D))}{P(D)} \\ &= \frac{\sum_i P(C_i \cap D)}{P(D)} = \sum_i P(C_i|D) \end{aligned}$$

(c)

$$\begin{aligned}
P(C \cap D) &= P(C \cap D \cap \Omega) \\
&= P(C \cap D \cap (\cup_i E_i)) \\
&= P(\cup_i (C \cap D \cap E_i)) \\
&= \sum_i P(C \cap D \cap E_i) \\
&= \sum_i P(E_i \cap D) P(C | E_i \cap D)
\end{aligned}$$

Divide both sides by $P(D)$ and we get

$$P(C|D) = \sum_i P(E_i|D) P(C|E_i \cap D).$$

(d) $P(A|C_i) = P(B|C_i)$ tells us that $P(A \cap C_i) = P(B \cap C_i)$.

$$\begin{aligned}
P(A \cap (\cup_i C_i)) &= P(\cup_i (A \cap \cup_i C_i)) \\
&= \sum_i P(A \cap \cup_i C_i) \\
&= \sum_i P(B \cap \cup_i C_i) \\
&= P(\cup_i (B \cap \cup_i C_i)) \\
&= P(B \cap (\cup_i C_i))
\end{aligned}$$

Similar to part (c), divide both sides by $P(\cup_i C_i)$, we get

$$P(A| \cup_i C_i) = P(B| \cup_i C_i).$$

Q4. Compute the distribution of $X + Y$ in the following cases:

(a) For $0 \le z \le n + m$,

$$\begin{aligned}
P(X + Y = z) &= \sum_{k=0}^{z} P(X = k) P(Y = z - k) \\
&= \sum_{k=0}^{z} \binom{n}{k} p^k (1-p)^{n-k} \binom{m}{z-k} p^{z-k} (1-p)^{m-z+k} \\
&= p^z (1-p)^{n+m-z} \sum_{k=0}^{z} \binom{n}{k} \binom{m}{z-k} \\
&= \binom{n+m}{z} p^z (1-p)^{n+m-z}.
\end{aligned}$$

In the above, we assume that $\binom{v}{r} = 0$ if $r > v$.
For $z < 0$ or $z > n + m$, we obviously have $P(X + Y = z) = 0$. Therefore, $X + Y$ is a binomial random variable with parameters $(n + m, p)$.

(b) For $n \geq 0$,

$$P(X+Y=n) = \sum_{k=0}^{n} P(X=k)P(Y=n-k) = \sum_{k=0}^{n} \frac{\lambda_1^k e^{-\lambda_1}}{k!} \cdot \frac{\lambda_2^{n-k} e^{-\lambda_2}}{(n-k)!}$$
$$= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda_1^k \lambda_2^{n-k}$$
$$= \frac{(\lambda_1+\lambda_2)^n e^{-(\lambda_1+\lambda_2)}}{n!}$$

For $n < 0$ we obviously have $P(X+Y=z) = 0$. Therefore, $X+Y$ is a Possion random variables with mean $\lambda_1 + \lambda_2$.

(c) The probaility distribution function of $Z = X+Y$ is

$$f_Z(z) = \int_{\mathbb{R}} f_X(z-y) f_Y(y) dy$$
$$= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(z-y-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} dy, \qquad z \in \mathbb{R}$$
$$= \int_{\mathbb{R}} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2} y^2 + \left(\frac{z-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right) y - \frac{(z-\mu_1)^2}{2\sigma_1^2} - \frac{\mu_2^2}{2\sigma_2^2}\right] dy$$
$$= \int_{\mathbb{R}} \exp\left[-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(y - \frac{(z-\mu_1)\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2\right] dy$$
$$\cdot \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[\frac{((z-\mu_1)\sigma_2^2+\mu_2\sigma_1^2)^2}{2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)} - \frac{(z-\mu_1)^2}{2\sigma_1^2} - \frac{\mu_2^2}{2\sigma_2^2}\right]$$
$$= \sqrt{\frac{2\pi\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}} \cdot \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[\frac{-(z-\mu_1)^2\sigma_1^2\sigma_2^2 + 2(z-\mu_1)\mu_2\sigma_1^2\sigma_2^2 - \mu_2^2\sigma_1^2\sigma_2^2}{2\sigma_1^2\sigma_2^2(\sigma_1^2+\sigma_2^2)}\right]$$
$$= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left[-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right].$$

Therefore, $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Q5.

Theorem 0.1 (The weak law of large numbers)**.** *Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $E[X_i] = \mu$. Then, for any $\epsilon > 0$,*

$$P\left\{\frac{X_1 + \cdots + X_n}{n} - \mu \geq \epsilon\right\} \to 0 \text{ as } n \to \infty.$$

Read materials on *Law of Large Number* and *Central Limit Theorem* in the book "*A First Course in Probability*" by Ross (Chapter 8), and write down the statements of both theorems.

Theorem 0.2 (The strong law of large numbers)**.** *Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite* *mean* $E[X_i] = \mu$. *Then, with probability* 1,

$$\frac{X_1 + X_2 + \dots X_n}{n} \to \mu \text{ as } n \to \infty.$$

Theorem 0.3 (The central limit theorem)**.** *Let* $X_1, X_2, \dots$ *be a sequence of independent and identically distributed random variables each having mean* μ *and variance* σ^2. *Then the distribution of*

$$\frac{X_1 + \dots X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the standard normal as $n \to \infty$. *That is, for* $-\infty < a < \infty$,

$$P\{\frac{X_1 + \dots X_n - n\mu}{\sigma\sqrt{n}} \leq a\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx \text{ as } n \to \infty.$$