Southern University of Science and Technology
Department of Mathematics

MA204: Mathematical Statistics

Tutorial 1: Examples/Solutions

A. Mutual Independency $\Rightarrow$ Pairwise Independency

Three events, $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ are mutually independent *if and only if* (iff)

(a) $\Pr(\mathbb{A}_1 \cap \mathbb{A}_2) = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2)$, i.e., $\mathbb{A}_1$ and $\mathbb{A}_2$ are independent;

(b) $\Pr(\mathbb{A}_2 \cap \mathbb{A}_3) = \Pr(\mathbb{A}_2) \times \Pr(\mathbb{A}_3)$, i.e., $\mathbb{A}_2$ and $\mathbb{A}_3$ are independent;

(c) $\Pr(\mathbb{A}_1 \cap \mathbb{A}_3) = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_3)$, i.e., $\mathbb{A}_1$ and $\mathbb{A}_3$ are independent;

(d) $\Pr(\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3) = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2) \times \Pr(\mathbb{A}_3)$.

Example T1.1: Give an example such that $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ are pairwise independent but not mutually independent.

Solution: Suppose a box contains 4 tickets labeled as {112, 121, 211, 222}. Let's choose one ticket at random, and consider the following three events:

$$\begin{aligned}
\mathbb{A}_1 &= \{1 \text{ occurring at the first place}\},\\
\mathbb{A}_2 &= \{1 \text{ occurring at the second place}\},\\
\mathbb{A}_3 &= \{1 \text{ occurring at the third place}\}.
\end{aligned}$$

So we obtain

$$\Pr(\mathbb{A}_1) = \frac{1}{2}, \quad \Pr(\mathbb{A}_2) = \frac{1}{2}, \quad \Pr(\mathbb{A}_3) = \frac{1}{2}.$$

Since

$$\mathbb{A}_1 \cap \mathbb{A}_2 = \{112\}, \quad \Pr(\mathbb{A}_1 \cap \mathbb{A}_2) = \tfrac{1}{4} = \Pr(\mathbb{A}_1)\Pr(\mathbb{A}_2),$$

$$\mathbb{A}_2 \cap \mathbb{A}_3 = \{121\}, \quad \Pr(\mathbb{A}_2 \cap \mathbb{A}_3) = \tfrac{1}{4} = \Pr(\mathbb{A}_2)\Pr(\mathbb{A}_3),$$

$$\mathbb{A}_1 \cap \mathbb{A}_3 = \{211\}, \quad \Pr(\mathbb{A}_1 \cap \mathbb{A}_3) = \tfrac{1}{4} = \Pr(\mathbb{A}_1)\Pr(\mathbb{A}_3),$$

we have the conclusion that $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ are pairwise independent. On the other hand, note that $\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3 = \emptyset$. then,

$$\Pr(\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3) = 0 \neq \frac{1}{8} = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2) \times \Pr(\mathbb{A}_3).$$

So $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ are not mutually independent. $\|$

Example T1.2: Give an example satisfying

$$\Pr(\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3) = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2) \times \Pr(\mathbb{A}_3)$$

but $\mathbb{A}_1$, $\mathbb{A}_2$ and $\mathbb{A}_3$ are not pairwise independent.

Solution: Toss two different standard dice. The sample space $\mathbb{S}$ of the outcomes consists of all the ordered pairs:

$$\mathbb{S} = \left\{ \begin{array}{cccc} (1,1), & (1,2), & \cdots, & (1,6) \\ (2,1), & (2,2), & \cdots, & (2,6) \\ (3,1), & (3,2), & \cdots, & (3,6) \\ (4,1), & (4,2), & \cdots, & (4,6) \\ (5,1), & (5,2), & \cdots, & (5,6) \\ (6,1), & (6,2), & \cdots, & (6,6) \end{array} \right\}.$$

Each point in $\mathbb{S}$ has a probability of $1/36$. Consider the following three events:

$$\mathbb{A}_1 = \{\text{first die shows 1 or 2 or 3}\},$$

$$\mathbb{A}_2 = \{\text{first die shows 3 or 4 or 6}\},$$

$$\mathbb{A}_3 = \{\text{sum of two faces is 9}\}.$$

So we have

$$\Pr(\mathbb{A}_1) = \frac{1}{2}, \quad \Pr(\mathbb{A}_2) = \frac{1}{2}, \quad \Pr(\mathbb{A}_3) = \Pr\{(3,6),(4,5),(5,4),(6,3)\} = \frac{1}{9}.$$

Note that $\mathbb{A}_1 \cap \mathbb{A}_2 = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$, so

$$\Pr(\mathbb{A}_1 \cap \mathbb{A}_2) = \frac{6}{36} = \frac{1}{6} \neq \frac{1}{4} = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2).$$

That is, $\mathbb{A}_1$, $\mathbb{A}_2$, and $\mathbb{A}_3$ are not pairwise independent. However, $\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3 = \{(3,6)\}$, we obtain

$$\Pr(\mathbb{A}_1 \cap \mathbb{A}_2 \cap \mathbb{A}_3) = \frac{1}{36} = \Pr(\mathbb{A}_1) \times \Pr(\mathbb{A}_2) \times \Pr(\mathbb{A}_3) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{9}. \quad \|$$

B. Expectation, Variance and Chebyshev's Inequality

B.1 Expectation and Variance

Let X be a discrete (or continuous) r.v. with pmf (or pdf) $f(x)$, and $g(x)$ be an arbitrary function. Then $g(X)$ is also a r.v. and the expectation of $g(X)$ is defined as:

$$E[g(X)] = \begin{cases} \displaystyle\sum_x g(x)f(x), & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{\infty} g(x)f(x)\ \mathrm{d}x, & \text{if } X \text{ is continuous.} \end{cases}$$

The *expectation* and *variance* of X are defined as:

$$\mu \mathrel{\hat{=}} E(X) \quad \text{and} \quad \sigma^2 \mathrel{\hat{=}} \mathrm{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2.$$

B.2 Chebyshev's Inequality

Let X be a r.v. and c be a positive constant, then

$$\Pr(|X-\mu| \geq c\sigma) \leq \frac{1}{c^2}.$$

Example T1.3: Let the pdf of X be given by

$$f(x) = \begin{cases} \dfrac{1}{2\sqrt{3}}, & \text{if } -\sqrt{3} < x < \sqrt{3}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Calculate $\Pr(|X| \geq \frac{3}{2})$

(b) Check the answer by the Chebyshev inequality.

Solution: (a) According to definition, we calculate

$$\begin{aligned} \Pr\left(|X| \geq \frac{3}{2}\right) &= \Pr\left(X \geq \frac{3}{2} \text{ or } X \leq -\frac{3}{2}\right) \\ &= 1 - \Pr\left(-\frac{3}{2} \leq X \leq \frac{3}{2}\right) \end{aligned}$$

$$
\begin{aligned}
&= 1 - \int_{-3/2}^{3/2} f(x)\ \mathrm{d}x \\
&= 1 - \int_{-3/2}^{3/2} \frac{1}{2\sqrt{3}}\ \mathrm{d}x \\
&= 1 - \frac{1}{2\sqrt{3}}\left[\frac{3}{2} - \left(-\frac{3}{2}\right)\right] \\
&= 1 - \frac{\sqrt{3}}{2} \approx 0.134.
\end{aligned}
$$

(b) The mean of variance of X are given by

$$
\begin{aligned}
\mu &= E(X) = \int_{-\infty}^{\infty} x\, f(x)\ \mathrm{d}x = \int_{-\sqrt{3}}^{\sqrt{3}} \frac{x}{2\sqrt{3}}\ \mathrm{d}x = 0, \\
\sigma^2 &= E(X^2) - \mu^2 = \int_{-\infty}^{\infty} x^2\, f(x)\ \mathrm{d}x - 0 = \int_{-\sqrt{3}}^{\sqrt{3}} \frac{x^2}{2\sqrt{3}}\ \mathrm{d}x = 1.
\end{aligned}
$$

We want to check if

$$
\Pr(|X-\mu| \geq c\sigma) = \Pr(|X| \geq c) \leq \frac{1}{c^2}
$$

for some positive constant c. In fact, from (a), we have $\Pr(|X| \geq 3/2) \approx 0.134$. For $c = 3/2$,

$$
\Pr\left(|X| \geq \frac{3}{2}\right) \approx 0.134 \leq \frac{1}{(\frac{3}{2})^2} \approx 0.44,
$$

so the Chebyshev inequality holds. ‖

C. Conditional Expectation and Conditional Variance

C.1 Conditional Expectation

Let X and Y be r.v.'s and $f(x \mid y)$ be the conditional pmf (or pdf) of X given $Y = y$. Then the conditional expectation of $g(X)$ given $Y = y$ is:

$$
E[g(X) \mid Y = y] = \begin{cases} \displaystyle\sum_x g(x) f(x \mid y), & \text{if } X \text{ is discrete,} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f(x \mid y)\ \mathrm{d}x, & \text{if } X \text{ is continuous.} \end{cases}
$$

Note that $E[g(X) \mid Y]$ is a function of the r.v. Y and we can similarly define the *conditional expectation* and *conditional variance* as in the unconditional case.

C.2 **Calculation Formulae of Expectation and Variance**

It can be shown that

$$E(X) = E[E(X \mid Y)] = \int_{-\infty}^{\infty} E(X \mid Y = y) f(y)\ \mathrm{d}y,$$
$$\mathrm{Var}(X) = E[\mathrm{Var}(X \mid Y)] + \mathrm{Var}[E(X \mid Y)].$$

Example T1.4: Suppose that the conditional pdf of (X, Y) given the r.v. Z is

$$f(x, y \mid z) = [z + (1-z)(x+y)] I_{(0,1)}(x) I_{(0,1)}(y),$$

for $0 \le z \le 2$, and the density of Z is $f(z) = \frac{1}{2} I_{(0,2)}(z)$, where $I_{\mathbb{A}}(x)$ denotes the indicator function, i.e., $I_{\mathbb{A}}(x) = 1$ if $x \in \mathbb{A}$ and $I_{\mathbb{A}}(x) = 0$ if $x \notin \mathbb{A}$.

(a) Find the expectation $E(X+Y)$.

(b) Determine whether X and Y are independent or not.

(c) Determine whether X and Z are independent or not.

Solution: (a) Note that $E(X+Y) = E[E(X+Y \mid Z)]$. We first calculate

$$\begin{aligned}
E(X+Y \mid Z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)[z+(1-z)(x+y)] I_{(0,1)}(x) I_{(0,1)}(y)\ \mathrm{d}x\ \mathrm{d}y \\
&= \int_0^1\int_0^1 (x+y)[z+(1-z)(x+y)]\ \mathrm{d}x\ \mathrm{d}y \\
&= \int_0^1\int_0^1 xz + yz + (1-z)x^2 + 2(1-z)xy + (1-z)y^2\ \mathrm{d}x\ \mathrm{d}y \\
&= \int_0^1 \left[\frac{x^2 z}{2} + xyz + \frac{(1-z)x^3}{3} + (1-z)x^2 y + (1-z)xy^2\right]\Bigg|_0^1\ \mathrm{d}y \\
&= \int_0^1 \frac{2+z}{6} + y + (1-z)y^2\ \mathrm{d}y \\
&= \left[\frac{(2+z)y}{6} + \frac{y^2}{2} + \frac{(1-z)y^3}{3}\right]\Bigg|_0^1 = \frac{7-z}{6},
\end{aligned}$$

so that

$$\begin{aligned}
E(X+Y) &= E[E(X+Y \mid Z)] = \int_{-\infty}^{\infty} E(X+Y \mid Z = z) f(z)\ \mathrm{d}z \\
&= \int_0^2 \frac{7-z}{6}\cdot\frac{1}{2}\ \mathrm{d}z = \left[\frac{7z}{12} - \frac{z^2}{24}\right]\Bigg|_0^2 = 1.
\end{aligned}$$

(b) Since

$$f(x,y,z) = f(x,y \mid z)f(z) = \frac{1}{2}[z + (1-z)(x+y)]I_{(0,1)}(x)I_{(0,1)}(y)I_{(0,2)}(z),$$

we have

$$\begin{aligned}
f(x,y) &= \int_{-\infty}^{\infty} f(x,y,z)\ \mathrm{d}z \\
&= \int_{-\infty}^{\infty} \frac{1}{2}[z + (1-z)(x+y)]I_{(0,1)}(x)I_{(0,1)}(y)I_{(0,2)}(z)\ \mathrm{d}z \\
&= \int_{0}^{2} \frac{1}{2}[z + (1-z)(x+y)]I_{(0,1)}(x)I_{(0,1)}(y)\ \mathrm{d}z \\
&= \frac{1}{2}I_{(0,1)}(x)I_{(0,1)}(y)\int_{0}^{2} x + y + (1-x-z)z\ \mathrm{d}z \\
&= \frac{1}{2}I_{(0,1)}(x)I_{(0,1)}(y)\left[(x+y)z + \frac{(1-x-y)z^2}{2}\right]\Bigg|_0^2 \\
&= I_{(0,1)}(x)I_{(0,1)}(y).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
f(x) &= \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}y = \int_{-\infty}^{\infty} I_{(0,1)}(x)I_{(0,1)}(y)\ \mathrm{d}y \\
&= \int_{0}^{1} I_{(0,1)}(x)\ \mathrm{d}y = I_{(0,1)}(x), \\
f(y) &= \int_{-\infty}^{\infty} f(x,y)\ \mathrm{d}x = \int_{-\infty}^{\infty} I_{(0,1)}(x)I_{(0,1)}(y)\ \mathrm{d}x \\
&= \int_{0}^{1} I_{(0,1)}(y)\ \mathrm{d}x = I_{(0,1)}(y).
\end{aligned}$$

Therefore, we obtain

$$f(x,y) = f(x)f(y),$$

i.e., X and Y are independent.

(c) Note that

$$\begin{aligned}
f(x,z) &= \int_{-\infty}^{\infty} f(x,y,z)\ \mathrm{d}y \\
&= \int_{-\infty}^{\infty} \frac{1}{2}[z + (1-z)(x+y)]I_{(0,1)}(x)I_{(0,1)}(y)I_{(0,2)}(z)\ \mathrm{d}y \\
&= \int_{0}^{1} \frac{1}{2}[z + (1-z)(x+y)]I_{(0,1)}(x)I_{(0,2)}(z)\ \mathrm{d}y
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} I_{(0,1)}(x) I_{(0,2)}(z) \int_0^1 z + (1-z)x + (1-z)y \ \mathrm{d}y \\
&= \frac{1}{2} I_{(0,1)}(x) I_{(0,2)}(z) \left[yz + (1-z)xy + \frac{(1-z)y^2}{2} \right] \Big|_0^1 \\
&= \frac{1+2x+z-2xz}{4} I_{(0,1)}(x) I_{(0,2)}(z) \\
&\neq f(x)f(z) = \frac{1}{2} I_{(0,1)}(x) I_{(0,2)}(z),
\end{aligned}$$

then, X and Z are not independent. $\|$

Example T1.5: Let $X_1, X_2 \overset{\text{iid}}{\sim} N(0,1)$, $U \sim U(0,1)$, and U be independent of (X_1, X_2). Define $Z = UX_1 + (1-U)X_2$.

(a) Find the conditional distribution of Z given $U = u$.

(b) Find $E(Z)$ and $\mathrm{Var}(Z)$.

(c) Find the distribution of Z.

Solution: (a) $Z|(U=u) = uX_1 + (1-u)X_2 \sim N(0, u^2 + (1-u)^2)$. Hence,

$$Z|U \sim N(0, U^2 + (1-U)^2)$$

so that $E(Z|U) = 0$ and $\mathrm{Var}(Z|U) = U^2 + (1-U)^2$.

(b) $E(Z) = E[E(Z|U)] = 0$.

$$\begin{aligned}
\mathrm{Var}(Z) &= E[\mathrm{Var}(Z|U)] + \mathrm{Var}[E(Z|U)] \\
&= E(U^2) + E(1-U)^2 + 0 \\
&= \int_0^1 u^2 \ \mathrm{d}u + \int_0^1 (1-u)^2 \ \mathrm{d}u \\
&= \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.
\end{aligned}$$

(c) Let $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ denote the cdf of $X \sim N(0,1)$. The cdf of Z is given by

$$\begin{aligned} F_Z(z) &= \Pr(Z \leq z) \\ &= \int_0^1 \Pr(Z \leq z | U = u) \cdot f_U(u) \text{ d}u \\ &= \int_0^1 \Pr\left(\frac{Z}{\sqrt{u^2 + (1-u)^2}} \leq \frac{z}{\sqrt{u^2 + (1-u)^2}} \middle| U = u \right) \text{d}u \\ &= \int_0^1 \left\{ \int_{-\infty}^{z/\sqrt{u^2+(1-u)^2}} \phi(x)\, dx \right\} \text{ d}u. \end{aligned}$$

$\|$