# AI 5030: PROBABILITY AND STOCHASTIC PROCESSES

INSTRUCTOR: DR. KARTHIK P. N.

## HOMEWORK 8

## TOPICS: EXPECTATIONS OF DISCRETE AND CONTINUOUS RANDOM VARIABLES, VARIANCE, COVARIANCE

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. All random variables appearing below are assumed to be defined with respect to $\mathscr{F}$.

1. For any $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer lesser than or equal to $x$. Thus, for instance, $\lfloor 3.5 \rfloor = 3$, $\lfloor -8.9 \rfloor = -9$, $\lfloor 2 \rfloor = 2$, and so on.
Suppose that $X \sim$ Exponential$(1)$. Determine the expected value of $Y = \lfloor X \rfloor$.

**Solution:** For any $k \in \mathbb{N} \cup \{0\}$, we have

$$\mathbb{P}(\{Y = k\}) = \mathbb{P}(\{k \leq X < k+1\}) = \mathbb{P}(\{k < X \leq k+1\}) = F_X(k+1) - F_X(k) = e^{-k} - e^{-(k+1)}.$$

We then have

$$\mathbb{E}[Y] = \sum_{k=0}^{\infty} k\, \mathbb{P}(\{Y = k\}) = \sum_{k=0}^{\infty} k(e^{-k} - e^{-(k+1)}) = (1 - e^{-1}) \sum_{k=1}^{\infty} k e^{-k} = (1 - e^{-1}) \left[\frac{e^{-1}}{1 - e^{-1}} + \frac{e^{-1}}{(1 - e^{-1})^2}\right] = \frac{2e - 1}{e - 1}.$$

2. Let $X$ be a non-negative and continuous random variable with PDF $f_X$ and CDF $F_X$. Show that

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(\{X > x\})\, \mathrm{d}x = \int_0^{\infty} (1 - F_X(x))\, \mathrm{d}x,$$

where the above integrals are usual Riemann integrals.

Hint: Write down the formula for expectation in terms of the PDF, and apply change of order of integration.

**Solution:** We have

$$\begin{aligned}
\int_0^{\infty} \mathbb{P}(\{X > x\})\, \mathrm{d}x &= \int_0^{\infty} \int_x^{\infty} f_X(t)\, \mathrm{d}t\, \mathrm{d}x \\
&\overset{(a)}{=} \int_0^{\infty} \int_0^{t} f_X(t)\, \mathrm{d}x\, \mathrm{d}t \\
&= \int_0^{\infty} t\, f_X(t)\, \mathrm{d}t \\
&= \mathbb{E}[X],
\end{aligned}$$

where $(a)$ above follows from the change of order of integration, and the last line follows from the formula for the expectation of a non-negative random variable in terms of its PDF.

3. Suppose that $X$ and $Y$ are jointly discrete random variables. The random variable $X$ takes values in $\{-1, 0, 1\}$ with uniform probabilities. Suppose that for each $x \in \{-1, 0, 1\}$,

$$p_{Y|X=x}(y) = \frac{1}{2}\, \mathbf{1}_{\{|y-x|=1\}}, \qquad y \in \mathbb{R}.$$

Compute $\mathbb{E}[Y]$.

**Solution:** Note that

$$p_{Y|X=x}(y) = \begin{cases} \frac{1}{2}, & y = x - 1, \\ \frac{1}{2}, & y = x + 1, \\ 0, & \text{otherwise.} \end{cases}$$

It follows then that the set of all possible values that $Y$ can take is $\{-2,-1,0,1,2\}$. Furthermore, we note that

$$\begin{aligned}
\mathbb{P}(\{Y=2\}) &= p_{Y|X=1}(2)\cdot\mathbb{P}(\{X=1\}) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6},\\
\mathbb{P}(\{Y=1\}) &= p_{Y|X=0}(1)\cdot\mathbb{P}(\{X=0\}) = \frac{1}{6},\\
\mathbb{P}(\{Y=0\}) &= p_{Y|X=1}(0)\cdot\mathbb{P}(\{X=1\}) + p_{Y|X=-1}(0)\cdot\mathbb{P}(\{X=-1\}) = \frac{1}{3},\\
\mathbb{P}(\{Y=-1\}) &= p_{Y|X=0}(-1)\cdot\mathbb{P}(\{X=0\}) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6},\\
\mathbb{P}(\{Y=-2\}) &= p_{Y|X=-1}(-2)\cdot\mathbb{P}(\{X=-1\}) = \frac{1}{6}.
\end{aligned}$$

Thus, we have

$$\mathbb{E}[Y] = 2\cdot\frac{1}{6} + 1\cdot\frac{1}{6} + 0\cdot\frac{1}{3} + (-1)\cdot\frac{1}{6} + (-2)\cdot\frac{1}{6} = 0.$$

4. Let $X_1, X_2, \ldots \overset{\text{i.i.d.}}{\sim}$ Exponential($\lambda$), and let $N \sim$ Geometric($p$) be independent of $\{X_1, X_2, \ldots\}$. Here, $\lambda > 0$ and $p \in (0,1)$ are fixed constants. Compute $\mathbb{E}\left[\sum_{i=1}^{N} X_i\right]$.

**Solution:** Note that for any $n \in \mathbb{N}$,

$$S_N\,\mathbf{1}_{\{N=n\}} = S_n\,\mathbf{1}_{\{N=n\}} = (X_1+\cdots+X_n)\,\mathbf{1}_{\{N=n\}}.$$

Because $N \sim$ Geometric($p$), we have $\mathbb{E}[N] = 1/p < +\infty$, which when combined with the fact that $N$ is a non-negative random variable yields $\mathbb{P}(\{N < +\infty\}) = 1$ (see homework 7, question 6(b)). Thus, we have

$$S_N = \sum_{n=1}^{\infty} (X_1+\cdots+X_n)\,\mathbf{1}_{\{N=n\}}.$$

For each $k \in \mathbb{N}$, let $Y_k = \sum_{n=1}^{k} (X_1+\cdots X_n)\,\mathbf{1}_{\{N=n\}}$. Observe that

$$\forall \omega \in \Omega, \qquad 0 \le Y_1(\omega) \le Y_2(\omega) \le \cdots,$$

$$\forall \omega \in \Omega, \qquad \lim_{k\to\infty} Y_k(\omega) = \lim_{k\to\infty} \sum_{n=1}^{k} (X_1(\omega)+\cdots+X_n(\omega))\,\mathbf{1}_{\{N=n\}}(\omega) = \sum_{n=1}^{\infty} (X_1(\omega)+\cdots+X_n(\omega)\,\mathbf{1}_{\{N=n\}}(\omega).$$

Using the monotone convergence theorem, we get

$$\begin{aligned}
\mathbb{E}[S_N] &= \lim_{k\to\infty} \mathbb{E}[Y_k]\\
&= \lim_{k\to\infty} \sum_{n=1}^{k} \mathbb{E}[(X_1+\cdots+X_n)\,\mathbf{1}_{\{N=n\}}]\\
&\overset{(*)}{=} \lim_{k\to\infty} \sum_{n=1}^{k} \mathbb{E}[(X_1+\cdots+X_n)]\cdot\mathbb{E}[\mathbf{1}_{\{N=n\}}]\\
&= \sum_{n=1}^{\infty} \frac{n}{\lambda}\,p_N(n)\\
&= \frac{1}{\lambda p},
\end{aligned}$$

where $(*)$ follows from the observation that $\mathbf{1}_{\{N=n\}}$ is independent of $(X_1+\cdots+X_n)$, which in turn is a consequence of the fact that $N \perp\!\!\!\perp \{X_1, X_2, \ldots\}$, and the last line follows by noting that $\mathbb{E}[N] = \sum_{n=1}^{\infty} n\,p_N(n) = 1/p$.

5. (a) Let $X$ and $Y$ be jointly continuous with the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} cx^2 + \dfrac{xy}{3}, & 0 \le x \le 1,\ 0 \le y \le 2, \\ 0, & \text{otherwise.} \end{cases}$$

   i. Find the constant $c$.
   ii. Are $X$ and $Y$ independent?
   iii. Calculate $\mathrm{Cov}(X, Y)$.

**Solution:** We solve each of the parts below.

   i. We have

$$1 = \int_0^1 \int_0^2 \left(cx^2 + \frac{xy}{3}\right)\,\mathrm{d}y\,\mathrm{d}x = \frac{2c}{3} + \frac{1}{3},$$

   from which it follows that $c = 1$.

   ii. For any $x \in [0, 1]$, we have

$$f_X(x) = \int_0^2 \left(x^2 + \frac{xy}{3}\right)\,\mathrm{d}y = 2x^2 + \frac{2x}{3},$$

   and for any $y \in [0, 2]$, we have

$$f_Y(y) = \int_0^1 \left(x^2 + \frac{xy}{3}\right)\,\mathrm{d}x = \frac{1}{3} + \frac{y}{6}.$$

   Clearly, $f_{X,Y}(1, 0) = 1 \neq \frac{8}{9} = f_X(1)\, f_Y(0)$, thereby establishing that $X \not\perp\!\!\!\perp Y$.

   iii. Note that

$$\mathbb{E}[XY] = \int_0^1 \int_0^2 \left(x^3 y + \frac{x^2 y^2}{3}\right)\,\mathrm{d}y\,\mathrm{d}x = \frac{43}{54},$$

$$\mathbb{E}[X] = \int_0^1 \left(2x^2 + \frac{2x}{3}\right)\,\mathrm{d}x = \frac{13}{18},$$

$$\mathbb{E}[X] = \int_0^2 \left(\frac{1}{3} + \frac{y}{6}\right)\,\mathrm{d}y = \frac{10}{9}.$$

   Thus, we have

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = -\frac{1}{162}$$

6. Let $X$ and $Y$ be independent random variables distributed uniformly on $[0, 1]$.
Let $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$. Calculate $\mathrm{Cov}(U, V)$.

**Solution:** Note that $U \cdot V = X \cdot Y$. Thus,

$$\mathbb{E}[UV] = \mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y] = \frac{1}{4},$$

where the penultimate equality follows because $X \perp\!\!\!\perp Y$. Furthermore, for any $u \in [0, 1]$,

$$\mathbb{P}(\{U > u\}) = \mathbb{P}(\{X > u\} \cap \{Y > u\}) = (1 - u)^2,$$

from which it follows that

$$\mathbb{E}[U] = \int_0^\infty \mathbb{P}(\{U > u\})\,\mathrm{d}u = \int_0^1 (1 - u)^2\,\mathrm{d}u = \int_0^1 u^2\,\mathrm{d}u = \frac{1}{3}.$$

Along similar lines, we note that for all $v \in [0, 1]$,

$$F_V(v) = \mathbb{P}(\{V \le v\}) = \mathbb{P}(\{X \le v\} \cap \{Y \le v\}) = v^2,$$

from which it follows that

$$\mathbb{E}[V] = \int_0^1 \mathbb{P}(\{V > v\})\,\mathsf{d}v = \int_0^1 (1 - v^2)\,\mathsf{d}v = 1 - \frac{1}{3} = \frac{2}{3}.$$

Combining the above results, it follows that

$$\mathsf{Cov}(U, V) = \mathbb{E}[UV] - \mathbb{E}[U]\,\mathbb{E}[V] = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}.$$

7. Let $X \sim \mathcal{N}(0, 1)$. Let $W$ be a discrete random variable independent of $X$ and having the PMF

$$\mathbb{P}(\{W = w\}) = \begin{cases} \frac{1}{2}, & w = \pm 1, \\ 0, & \text{otherwise.} \end{cases}$$

Define a new random variable $Y$ as $Y = WX$.

(a) Show that $Y \sim \mathcal{N}(0, 1)$.
(b) Show that $X$ and $Y$ are uncorrelated, but not independent.
(c) A friend of yours comes to you and claims that $Z = X + Y$ is Gaussian distributed. Is your friend's claim correct?

**Solution:** We provide the solution to each part below.

(a) For any $y \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{P}(\{Y \le y\}) &= \mathbb{P}(\{Y \le y\} \cap \{W = 1\}) + \mathbb{P}(\{Y \le y\} \cap \{W = -1\}) \\ &= \mathbb{P}(\{X \le y\} \cap \{W = 1\}) + \mathbb{P}(\{-X \le y\} \cap \{W = -1\}) \\ &\overset{(a)}{=} \mathbb{P}(\{X \le y\}) \cdot \mathbb{P}(\{W = 1\}) + \mathbb{P}(\{-X \le y\}) \cdot \mathbb{P}(\{W = -1\}) \\ &\overset{(b)}{=} \mathbb{P}(\{X \le y\}) \cdot \mathbb{P}(\{W = 1\}) + \mathbb{P}(\{X \le y\}) \cdot \mathbb{P}(\{W = -1\}) \\ &= \mathbb{P}(\{X \le y\}), \end{aligned}$$

where $(a)$ above follows because $W \perp\!\!\!\perp X$, and $(b)$ above follows by noting that $X$ and $-X$ have identical CDFs (as the PDF of $X$ is symmetric about the origin). Thus, it follows that $Y \sim \mathcal{N}(0, 1)$.

(b) Note that

$$\mathbb{E}[XY] = \mathbb{E}[WX^2] = \mathbb{E}[W]\,\mathbb{E}[X^2] = 0,$$

where the penultimate equality follows because $W \perp\!\!\!\perp X$ (therefore $W \perp\!\!\!\perp X^2$), and the last equality follows because $\mathbb{E}[W] = 0$. Furthermore, note that $\mathbb{E}[Y] = \mathbb{E}[X] = 0$. Therefore, it follows that

$$\mathsf{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0.$$

To show that $X \not\perp\!\!\!\perp Y$, observe that for any given $x \in \mathbb{R}$, conditioned on the event $\{X = x\}$, the random variable $Y$ is discrete and takes values $\pm x$ with equal probability, i.e.,

$$p_{Y|X=x}(y) = \frac{1}{2}\,\mathbf{1}_{\{x\}}(y) + \frac{1}{2}\,\mathbf{1}_{\{-x\}}(y), \qquad y \in \mathbb{R}.$$

Clearly, the conditional and unconditional distributions of $Y$ are not the same, thus proving that $X \not\perp\!\!\!\perp Y$.

(c) We have

$$\mathbb{P}(\{Z = 0\}) = \mathbb{P}(\{Y = -X\}) = \mathbb{P}(\{W = -1\}) = \frac{1}{2},$$

thus proving that $Z$ is not Gaussian distributed.

8. Fix $n \in \mathbb{N}, n \ge 2$.
Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed with finite mean $\mu$ and variance $\sigma^2$.
Define the sample mean $M_n$ and sample variance $V_n$ as the random variables

$$M_n := \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad V_n := \frac{1}{n-1}\sum_{i=1}^{n} (X_i - M_n)^2.$$

(a) Show that $\mathbb{E}[M_n] = \mu$.

(b) Show that Var$(M_n) = \frac{\sigma^2}{n}$.

(c) Show that $\mathbb{E}[V_n] = \sigma^2$ (the factor $(n-1)$ in the denominator of $V_n$ is precisely to ensure that the mean of $V_n$ is equal to $\sigma^2$).

**Solution:** We provide the solution to each part below.

(a) Using the linearity of expectations, we have

$$\mathbb{E}[M_n] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mu.$$

(b) We have

$$\text{Var}(M_n) \overset{(*)}{=} \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) \overset{(a)}{=} \frac{\sigma^2}{n},$$

where $(*)$ follows from the observation that Var$(aY) = a^2\,$Var$(Y)$, and $(a)$ follows by noting that $X_1, \ldots, X_n$ are independent, and thus the variance of their sum is the sum of their variances.

(c) We have

$$\begin{aligned}\mathbb{E}[V_n] &= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[(X_i - M_n)^2] \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[((X_i-\mu) - (M_n-\mu))^2] \\ &= \frac{1}{n-1}\sum_{i=1}^{n}\Big[\mathbb{E}[(X_i-\mu)^2] + \mathbb{E}[(M_n-\mu)^2] - 2\,\mathbb{E}[(X_i-\mu)(M_n-\mu)]\Big].\end{aligned}$$

We now note that $\mathbb{E}[(X_i-\mu)^2] = \text{Var}(X_i) = \sigma^2$, $\mathbb{E}[M_n] = \mu$,

$$\mathbb{E}[(M_n-\mu)^2] = \text{Var}(M_n) = \frac{\sigma^2}{n},$$

and for each $i \in \{1, \ldots, n\}$,

$$\mathbb{E}[(X_i-\mu)(M_n-\mu)] = \frac{1}{n}\,\mathbb{E}[(X_i-\mu)^2] + \frac{1}{n}\sum_{j\neq i}\mathbb{E}[(X_i-\mu)(X_j-\mu)] \overset{(**)}{=} \mathbb{E}[(X_i-\mu)^2] = \frac{\sigma^2}{n},$$

where $(**)$ follows by noting that $\mathbb{E}[(X_i-\mu)(X_j-\mu)] = \mathbb{E}[(X_i-\mu)]\cdot\mathbb{E}[(X_j-\mu)] = 0$ (here, we use the fact that $X_i \perp\!\!\!\perp X_j$ for all $j \neq i$). Combining the above results, we get

$$\mathbb{E}[V_n] = \frac{1}{n-1}\sum_{i=1}^{n}\left[\sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n}\right] = \frac{1}{n-1}\sum_{i=1}^{n}\frac{(n-1)\,\sigma^2}{n} = \sigma^2.$$

9. Suppose that $X, Y$, and $Z$ are three random variables defined with respect to $\mathscr{F}$.
Let the means of $Y$ and $Z$ be $\mu_Y$ and $\mu_Z$ respectively. Show that

$$\mathbb{E}[\max\{X,\mu_Y\} - \max\{X,\mu_Z\}] \leq |\mu_Y - \mu_Z| \cdot \mathbb{P}\bigg(\Big\{X \in \big[\min\{\mu_Y,\mu_Z\}, \max\{\mu_Y,\mu_Z\}\big]\Big\}\bigg).$$

Hint: Consider the cases $\mu_Y < \mu_Z$ and $\mu_Y \geq \mu_Z$ separately.
For each case, break down the sample space into events of the form $\{X < \mu_Y\}, \{\mu_Y \leq X \leq \mu_Z\}, \{X > \mu_Z\}$.
On each of these events, upper bound the mean value of $\max\{X,\mu_Y\} - \max\{X,\mu_Z\}$.

**Solution:** Let $W = \max\{X, \mu_Y\} - \max\{X, \mu_Z\}$. Then, we have

$$\begin{aligned}
&W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}} = W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{X<\mu_Y\}} + W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}} + W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{X>\mu_Z\}}\\
&= (\mu_Y-\mu_Z)\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{X<\mu_Y\}} + (X-\mu_Z)\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}} + (X-X)\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{X>\mu_Z\}}\\
&\overset{(a)}{\le} 0\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{X<\mu_Y\}} + |X-\mu_Z|\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}}\\
&\overset{(b)}{\le} |\mu_Y-\mu_Z|\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}}\\
&= |\mu_Y-\mu_Z|\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}\,\mathbf{1}_{\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\}}\\
&\le |\mu_Y-\mu_Z|\,\mathbf{1}_{\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\}}\,\mathbf{1}_{\{\mu_Y<\mu_Z\}},
\end{aligned}$$

where $(a)$ follows by noting that $(\mu_Y-\mu_Z)\,\mathbf{1}_{\{\mu_Y<\mu_Z\}} < 0$ and $X-\mu_Z \le |X-\mu_Z|$, and $(b)$ follows by noting that $|X-\mu_Z|\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}} \le |\mu_Y-\mu_Z|\,\mathbf{1}_{\{\mu_Y\le X\le\mu_Z\}}$. We then have

$$\mathbb{E}[W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}] \le |\mu_Y-\mu_Z|\cdot\mathbb{P}(\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\})\cdot\mathbf{1}_{\{\mu_Y<\mu_Z\}}.$$

Along similar lines, interchanging the roles of $\mu_Y$ and $\mu_Z$, we get

$$\mathbb{E}[W\,\mathbf{1}_{\{\mu_Y\ge\mu_Z\}}] \le |\mu_Y-\mu_Z|\cdot\mathbb{P}(\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\})\cdot\mathbf{1}_{\{\mu_Y\ge\mu_Z\}}.$$

Adding the results obtained above, we get

$$\begin{aligned}
\mathbb{E}[W] &= \mathbb{E}[W\,\mathbf{1}_{\{\mu_Y<\mu_Z\}}] + \mathbb{E}[W\,\mathbf{1}_{\{\mu_Y\ge\mu_Z\}}]\\
&\le |\mu_Y-\mu_Z|\cdot\mathbb{P}(\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\})\cdot\mathbf{1}_{\{\mu_Y<\mu_Z\}}\\
&\qquad\qquad + |\mu_Y-\mu_Z|\cdot\mathbb{P}(\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\})\cdot\mathbf{1}_{\{\mu_Y\ge\mu_Z\}}\\
&= |\mu_Y-\mu_Z|\cdot\mathbb{P}(\{\min\{\mu_Y,\mu_Z\}\le X\le\max\{\mu_Y,\mu_Z\}\}).
\end{aligned}$$