# MAST20004 Probability
# Semester 1, 2021
# Assignment 4: Questions

Due 3 pm, Friday 21 May 2021

| **Name:** James La Fontaine |
|---|
| **Student ID:** 1079860 |

**Important instructions:**

(1) This assignment contains 5 questions, **two** of which will be randomly selected to be marked. Each marked question is worth 10 points and each unmarked question with substantial working is worth 1 point.

(2) To complete this assignment, you need to write your solutions into the blank answer spaces following each question in this assignment PDF.

- If you have a printer (or can access one), then you must print out the assignment template and handwrite your solutions into the answer spaces.
- If you do not have a printer but you can figure out how to annotate a PDF using an iPad/Android tablet/Graphics tablet or using Adobe Acrobat, then annotate your answers directly onto the assignment PDF and save a copy for submission.

Failing both of these methods, you may handwrite your answers as normal on blank paper and then scan for submission (but note that you will thereby miss valuable practice for the exam process). In that case, however, your document should have the same length as the assignment template otherwise Gradescope will reject your submission. So you will need to add as many blank pages as necessary to reach that criterion.

Scan your assignment to a PDF file using your mobile phone (we recommend Cam - Scanner App), then upload by going to the Assignments menu on Canvas and submit the PDF to the **GradeScope tool** by first selecting your PDF file and then clicking on 'Upload PDF'.

Note that here you do not need to submit any Matlab code with your assignment.

(3) A poor presentation penalty of 10% of the total available marks will apply unless your submitted assignment meets all of the following requirements:

- it is a single pdf with all pages in correct template order and the correct way up, and with any blank pages with additional working added only at the end of the template pages;
- has all pages clearly readable;
- has all pages cropped to the A4 borders of the original page and is imaged from directly above to avoid excessive 'keystoning'.

These requirements are easy to meet if you use a scanning app on your phone and take some care with your submission - please review it before submitting to double check you have satisfied all of the above requirements.

(4) Late submission within 20 hours after the deadline will be penalised by 5% of the total available marks for every hour or part thereof after the deadline. After that, the Gradescope submission channel will be closed, and your submission will no longer be accepted. You are strongly encouraged to submit the assignment a few days before the deadline just in case of unexpected technical issues. If you are facing a rather exceptional/extreme situation that prevents you from submitting on time, please contact the tutor coordinator **Robert Maillardet** with formal proofs such as medical certificate.

(5) Working and reasoning must be given to obtain full credit. Clarity, neatness, and style count.

**Q1.** Let $X_1 \stackrel{d}{=} \text{Pn}(1)$ and $X_2 \stackrel{d}{=} \text{Pn}(4)$ be two independent Poisson random variables. Let $Y = X_1 + X_2$; it is shown on lecture slide 405 that $Y \stackrel{d}{=} \text{Pn}(5)$.

(a) Given $Y = n$, $n \geq 0$, what are the possible values of $X_1$?

$0 \leq X_1 \leq n$

(b) Calculate the conditional distribution of $X_1$ given $Y = n$ for $n \geq 0$. Specify the distribution and its parameters.

$$p_{X_1,Y}(x,n) = p_{X_1,X_2}(x,n-x)$$

$$= p_{X_1}(x) \cdot p_{X_2}(n-x)$$

$$= \frac{1^x e^{-1}}{x!} \cdot \frac{4^{n-x} e^{-4}}{(n-x)!} = \frac{1^x 4^{n-x} e^{-5}}{x!\,(n-x)!}$$

$$p_{X_1|Y}(x,n) = \frac{1^x 4^{n-x} e^{-5}}{x!\,(n-x)!} \cdot \frac{n!}{5^n e^{-5}}$$

$$= \binom{n}{x} \frac{1^x 4^{n-x}}{5^n} = \binom{n}{x} \frac{1^x 4^{n-x} 5^{-x}}{5^n 5^{-x}}$$

$$= \binom{n}{x} \frac{1^x 4^{n-x}}{5^x 5^{n-x}}$$

$$= \binom{n}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{n-x}$$

$$\Rightarrow X_1 | Y \stackrel{d}{=} Bi\left(n, \frac{1}{5}\right)$$

(c) Compute the conditional expectation $\mathbb{E}(X_1|Y = n)$, the conditional variance $\mathbb{V}(X_1|Y = n)$ for $n \geq 0$, and express the conditional mean $\mathbb{E}(X_1|Y)$ and the conditional variance $\mathbb{V}(X_1|Y)$ as transformations of the random variable $Y$.

$$E[X_1|Y=n] = \sum_x x\, p_{X_1|Y}(x|n) = \sum_x x \binom{n}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{n-x}$$

$$= np$$

$$= \frac{n}{5}$$

$$V[X_1|Y=n] = np(1-p)$$

$$= \frac{4n}{25}$$

$$E[X_1|Y] = \frac{Y}{5}$$

$$V[X_1|Y] = \frac{4Y}{25}$$

(d) Evaluate $\mathbb{E}(X_1)$ and $\mathrm{V}(X_1)$ using $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}(X|Y)]$ and $\mathrm{V}[X] = \mathbb{E}[V(X|Y)] + \mathrm{V}[\mathbb{E}(X|Y)]$ with $X = X_1$ and $Y = X_1 + X_2$.

$$E[X_1] = E[E[X_1|Y]]$$
$$= E\left[\frac{Y}{5}\right]$$
$$= \sum_{y=0}^{\infty} \frac{y}{5} \frac{5^y e^{-5}}{y!} = \frac{1}{5} e^{-5} \sum_{y=0}^{\infty} y \frac{5^y}{y(y-1)!}$$
$$= 5 \cdot \frac{1}{5} \cdot e^{-5} \sum_{y=0}^{\infty} \frac{5^{y-1}}{(y-1)!} = e^{-5} \cdot e^{5} = 1$$

$$V[X_1] = E\left[\frac{4Y}{25}\right] + V\left[\frac{Y}{5}\right]$$
$$= E\left[\frac{4Y}{25}\right] + E\left[\frac{1}{25}Y^2\right] - E\left[\frac{Y}{5}\right]^2$$
$$= \frac{4}{5} E\left[\frac{Y}{5}\right] + \frac{1}{25} E[Y^2] - 1$$
$$= \frac{4}{5} + \frac{1}{25} \left[E[Y(Y-1)] + E[Y]\right] - 1$$
$$= \frac{4}{5} + \frac{1}{25} \left[\left(\sum_{y=0}^{\infty} y(y-1) \frac{5^y e^{-5}}{y!}\right) + 5\right] - 1$$
$$= \frac{4}{5} + \frac{1}{25} \left[\left(25 e^{-5} \sum_{y=0}^{\infty} \frac{5^{y-2}}{(y-2)!}\right) + 5\right] - 1$$
$$= \frac{4}{5} + \frac{30}{25} - 1 = 1$$

**Q2** (a) Prove Markov's inequality for nonnegative continuous random variables; that is, show that if $X$ is a nonnegative continuous random variable, then for any $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

*Hint: You may wish to start with* $\mathbb{E}[X] = \int_0^a x f(x)dx + \int_a^\infty x f(x)dx$, *where $f$ is the probability density function of $X$.*

$X \geq 0$, continuous, $a > 0$

$$E[X] = \int_0^a x f(x)\, dx + \int_a^\infty x f(x)\, dx$$

$$\geq \int_a^\infty x f(x)\, dx$$

$$\geq \int_a^\infty a f(x)\, dx$$

$$= a \int_a^\infty f(x)\, dx = a \left[ P_r(X \leq \infty) - P_r(X \leq a) \right]$$

$$= a P_r(X \geq a)$$

$$\Rightarrow E[X] \geq a P_r(X \geq a)$$

$$P_r(X \geq a) \leq \frac{E[X]}{a}$$

$\square$

(b) Prove Chebyshev's inequality for continuous random variables using Markov's inequality.

$$P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

let $\varepsilon = k\sigma > 0$

$$P[|X-\mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$$

$$P(|X-\mu| \geq \varepsilon)$$

$$= P((X-\mu)^2 \geq \varepsilon^2) \leq \frac{E((X-\mu)^2)}{\varepsilon^2}$$ by Markov's Inequality

$$\Rightarrow P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

$$\Rightarrow P\left(\left|\frac{X-\mu}{\sigma}\right| \geq k\right) \leq \frac{1}{k^2}$$

$\square$

**Q3** (a) Let $X$ be a nonnegative integer valued random variable with probability mass function $p_X(k)$. Derive an expression for $\mathbb{P}(X \text{ is odd})$ in terms of its probability generating function $P_X(z)$ defined on some domain $C \subset \mathbb{R}$.

$$P_x(z) = \sum_{k=0}^{\infty} p_x(k) \cdot z^k$$

$$1-(-1)^k = \begin{cases} 2 \text{ if } k \text{ is odd} \\ 0 \text{ if } k \text{ is even} \end{cases}$$

$$\Rightarrow P(X \text{ is odd}) = \frac{1}{2}\left(\sum_{k=0}^{\infty} p_x(k) \cdot 1 - \sum_{k=0}^{\infty} p_x(k)(-1)^k\right)$$

$$= \frac{1}{2}\left(P_x(1) - P_x(-1)\right)$$

$$= \frac{1}{2}\left(\sum_{k=0}^{\infty}(1-(-1)^k)\, p_x(k)\right)$$

(b) Let $X \stackrel{d}{=} \text{Pn}(\lambda)$ and $Y \stackrel{d}{=} \text{Nb}(r, p)$ be two independent random variables. Let $W = X + Y$.

(i) Write down the probability generating function for $W$, $P_W(z)$.

$$P_W(z) = P_X(z) \cdot P_Y(z)$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x z^x}{x!} \cdot \sum_{y=0}^{\infty} \binom{-r}{y} p^r (p-1)^y z^y$$

$$= e^{-\lambda} e^{\lambda z} \cdot p^r \sum_{y=0}^{\infty} \binom{-r}{y} ((p-1)z)^y$$

$$= e^{-\lambda(1-z)} \cdot p^r (1-(1-p)z)^{-r}$$

(ii) Which values of the parameters $\lambda$, $r$, and $p$ will ensure that $W$ is more likely to take on even values than odd values? Justify your answer.

$P_W(z) = e^{-\lambda(1-z)} \cdot p^r (1-(1-p)z)^{-r}$

$P(W \text{ is odd}) = 1 - P(W \text{ is even})$

$P(W \text{ is odd}) < \frac{1}{2}$

$\frac{1}{2}(1 - P_W(-1)) < \frac{1}{2}$

$1 - P_W(-1) < 1$

$-P_W(-1) < 0$

$P_W(-1) > 0$

$e^{-2\lambda} p^r (1+(1-p))^{-r} > 0$

$e^{-2\lambda} > 0 \quad \forall \lambda$

$p^r > 0 \quad \forall r, \ p > 0$

$(1+(1-p))^{-r} > 0 \quad \forall r, \ p \in (0,1]$

$\Rightarrow$ W is more likely to take on even values for $\lambda \geq 0, r > 0, p \in (0,1]$

(iii) In terms of $\lambda$, $r$, and $p$, evaluate $p_W(1) = P(W = 1)$ in two different ways: one using the probability generating function, and the other using the convolution formula.

1) $p_W(1) = \frac{P'_W(0)}{1!} = P'_W(0)$

$$\frac{d}{dz}\left(e^{-\lambda(1-z)} p^r (1-(1-p)z)^{-r}\right)\Big|_{z=0}$$

$$= e^{-\lambda+\lambda z} r(1-p)((p-1)z+1)^{-r-1} p^r$$
$$+ \lambda e^{-\lambda+\lambda z} p^r (1-(1-p)z)^{-r} \Big|_{z=0}$$

$$= e^{-\lambda} \cdot -(p-1)r \cdot p^r + \lambda e^{-\lambda} p^r$$

$$= p^r e^{-\lambda}(\lambda - (p-1)r)$$

2) $p_W(1) = \sum_{x=0}^{1} p_Y(1-x)\, p_X(x)$

$$= p_Y(1)\, p_X(0) + p_Y(0)\, p_X(1)$$
$$= \binom{-r}{1} p^r (p-1)^1 \cdot \frac{e^{-\lambda}\lambda^0}{0!}$$
$$+ \binom{-r}{0} p^r (p-1)^0 \cdot \frac{e^{-\lambda}\lambda^1}{1!}$$
$$= -r p^r (p-1) e^{-\lambda} + p^r e^{-\lambda} \lambda$$
$$= p^r e^{-\lambda}(\lambda - (p-1)r)$$

**Q4** (a) Let $X_1, \cdots, X_n$ be independent exponential random variables with parameter $\alpha$. Compute the moment generating function of $Y = X_1 + \cdots + X_n$, and hence show that $Y \stackrel{d}{=} \gamma(n, \alpha)$.

$$M_X(t) = \frac{\alpha}{\alpha - t}$$

$$M_Y(t) = \left(\frac{\alpha}{\alpha - t}\right)^n$$

let $Z \stackrel{d}{=} \gamma(n, \alpha)$

$$M_Z(t) = \int_0^\infty \frac{e^{xt} \alpha^n e^{-\alpha x} x^{n-1} dx}{\Gamma(n)} = \alpha^n \int_0^\infty \frac{e^{-(\alpha - t)x} x^{n-1}}{\Gamma(n)} dx$$

$$= \frac{\alpha^n}{(\alpha - t)^n} \int_0^\infty \frac{e^{-u} u^{n-1}}{\Gamma(n)} du$$

$$= \left(\frac{\alpha}{\alpha - t}\right)^n = M_Y(t)$$

$$\Rightarrow Y \stackrel{d}{=} \gamma(n, \alpha)$$

(b) Let $X$ be a standard normal random variable. Show that the moment generating function of $X^2$ is $M_{X^2}(t) = (1-2t)^{-1/2}$, $t < 1/2$ (justify every step of your reasoning).

$X \stackrel{d}{=} N(0,1)$

$$E[e^{tX^2}] = M_{X^2}(t) = \int_{-\infty}^{\infty} \frac{e^{tx^2} e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2\left(\frac{1}{2}-t\right)} \, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}(1-2t)} \, dx$$

$u = x\sqrt{1-2t}$

$dx = \frac{1}{\sqrt{1-2t}} \, du$

$$\Rightarrow \frac{1}{\sqrt{1-2t}} \underbrace{\int_{-\infty}^{\infty} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \, du}_{=1}$$ (pdf of standard normal)

$\Rightarrow M_{X^2}(t) = (1-2t)^{-1/2}, \quad t < \frac{1}{2}$

(c) Let $X_1, \cdots, X_n$ be independent standard normal random variables. Compute the moment generating function of $X_1^2 + \cdots + X_n^2$, and hence identify its distribution and specify its parameters.

Let $Y = X_1^2 + \ldots + X_n^2$

$M_{X^2}(t) = (1-2t)^{-1/2}$

$M_Y(t) = (M_{X^2}(t))^n$

$= (1-2t)^{-n/2}$

$= \frac{1}{(1-2t)^{n/2}}$

from (a)

$Y \stackrel{d}{=} \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$

**Q5** The amount of leaves consumed by a koala per day follows some distribution with mean 700 grams and standard deviation 100 grams. Assume that the leaves demand of each koala is independent. Suppose that there are 65 koalas in a koala conservation reserve, and the keepers of the reserve order 46 kilos of fresh leaves every day.

(a) On a particular day, what is (an approximation of) the probability that the leaves supplied by the reserve will not meet the demand of the koalas?

$\mu = 700, \sigma = 100$

$D_i$ = demand of each koala

$X = D_1 + \ldots + D_{65} \overset{d}{\approx} N(65 \cdot 700, 65 \cdot 100^2)$

$P(X > 46000) = 1 - P(X \leq 46000)$

$= 1 - P\left(\frac{X - \mu}{\sigma} \leq \frac{46000 - \mu}{\sigma}\right)$

$= 1 - P(Z \leq 0.62)$

$= 1 - 0.7324$

$= 0.2676$

(b) Suppose that the amount of leaves consumed by the koalas is independent from day to day. What is (an approximation of) the probability the reserve will not meet the leaves demand of the koalas for at least 1 of the next 5 days?

$X$ = days with leaves demand not met $\overset{d}{\approx}$ $Bi(5, 0.2676)$

$$P(X \geq 1) = 1 - P(X = 0)$$

$$= 1 - \binom{5}{0} 0.2676^0 (1 - 0.2676)^5$$

$$= 1 - 1 \cdot 1 \cdot (0.7324)^5$$

$$= 1 - 0.2107$$

$$= 0.7893$$

(c) During the next year, what is (an approximation of) the probability that the reserve will not meet the leaves demand of the koalas for more than 90 days?

$X$ = # days with leaves demand not met $\overset{d}{\approx} Bi(365, 0.2626)$

$\Rightarrow X \overset{d}{\approx} N(365 \cdot 0.2626, 365 \cdot 0.2676 \cdot 0.7324)$

$\Rightarrow P(X > 90) = P(Z > \frac{90-\mu}{\sigma})$

$= P(Z > -0.1073)$

$= 1 - P(Z \leq -0.11)$

$= 1 - 0.4562$

$= 0.5438$