[theorem]Question

Theorem ??

Name: Sherwin Adrien Tiu Student Number: 35443258

Section: MATH 302 102

HW₆

1. Suppose that X has moment generating function

$$M_X(t) = \frac{1}{4}e^{-3t} + \frac{1}{2} + \frac{1}{4}e^t.$$

- (a) Find the mean and variance of X by differentiating the m.g.f. above.
- (b) Find the p.m.f. of X. Use your expression for the p.m.f. to check your answers from part (a).

(a)
$$M_X'(t) = -\frac{3}{4}e^{-3t} + \frac{1}{4}e^t$$

$$M_x''(t) = \frac{9}{4}e^{-3t} + \frac{1}{4}e^t$$

$$\mu = \mathbb{E}(x) = M_x'(0) = \frac{-3}{4} + \frac{1}{4} = -\frac{1}{2}$$

$$\sigma^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = M_x''(0) - (M_x'(0))^2 = (\frac{9}{4} + \frac{1}{4}) - (-\frac{1}{2})^2 = \frac{9}{4}$$

(b)
$$M'_x(t) = \sum_k e^{tk} \mathbb{P}(X = k)$$

$$X = -3 \to \mathbb{P}(-3) = \frac{1}{4}$$

$$X = 0 \to \mathbb{P}(0) = \frac{1}{2}$$

$$X = 1 \to \mathbb{P}(1) = \frac{1}{4}$$

$$\mathbb{E}(X) = -3(\frac{1}{4}) + 0(\frac{1}{2}) + 1(\frac{1}{4}) = -\frac{1}{2}$$

$$\mathbb{E}(X^2) = -3^2(\frac{1}{4}) + 0^2(\frac{1}{2}) + 1^2(\frac{1}{4}) = \frac{10}{4}$$

$$\sigma^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{9}{4}$$

2. You have two dice, one with three sides labeled 0, 1, 2 (perhaps this problem takes place in a two-dimensional world) and one with 4 sides, labeled 0, 1, 2, 3. Let X_1 be the outcome of rolling the first die, and X_2 the outcome of rolling the second. Denote the joint p.m.f. of X_1 and X_2 by p_{X_1,X_2} . That is, $p_{X_1,X_2}(m,n) := \mathbb{P}[X_1 = m, X_2 = n]$ for all $(m,n) \in \{0,1,2\} \times \{0,1,2,3\}$. The rolls are independent.

- (a) What is the joint p.m.f. of (X_1, X_2) ? $p(x, y) = (\frac{1}{3})(\frac{1}{4}) = \frac{1}{12}; x \in [0, 1, 2, y] \in [0, 1, 2, 3]$
- (b) Let $Y_1 = X_1 \cdot X_2$ and $Y_2 = \max\{X_1, X_2\}$. Make a table for the joint p.m.f. of (Y_1, Y_2) .

y_1y_2	0	1	2	3	p_{y_i}
0	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	p_{y_i} $\frac{3}{6}$
1	0	$\frac{\frac{\overline{6}}{1}}{12}$	0	0	$\frac{1}{12}$
2	0	0	$\frac{1}{6}$	0	$\frac{1}{6}$
3	0	0	0	$\frac{1}{12}$	1
4	0	0	$\frac{1}{12}$	0	$\frac{\overline{12}}{\overline{12}}$
5	0	0	0	0	0
6	0	0	0	$\frac{1}{12}$	$\frac{1}{12}$
p_{y_2}	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{5}{12}$	$\frac{3}{12}$	

(c) Are Y_1 and Y_2 independent?

$$\mathbb{P}(Y_1, Y_2) \stackrel{?}{=} \mathbb{P}(Y_1) \mathbb{P}(Y_2)$$

Take
$$Y_1 = 2, Y_2 = 1$$

$$\mathbb{P}(Y_1, Y_2) = 0$$

$$\mathbb{P}(Y_1)\mathbb{P}(Y_2) = \frac{2}{12} \times \frac{4}{12} \neq 0$$

Not Independent.

3. (Not marked) In this guided exercise you are asked to prove the version of the law of unconscious statistician from Section 6.1, and then to use it in order to give a short and easy proof of (part (b)) linearity of expectation and (part (c)) of the fact that expectation of a product of independent random variables is the product of the expectations:

Let X_1 and X_2 be two discrete random variables. In parts (b) and (c) suppose that either $\mathbb{P}[X \geq 0, Y \geq 0] = 1$ or that $\mathbb{E}[|X|]$ and $\mathbb{E}[|Y|]$ are both finite (or both). Prove the following claims from lecture:

(a) If $g: \mathbb{R}^2 \to \mathbb{R}$ is a function, then

$$\mathbb{E}[g(X_1, X_2)] = \sum_{k_1, k_2} g(k_1, k_2) \cdot \mathbb{P}(X_1 = k_1, X_2 = k_2).$$

Hint: Let $A_{\ell} := \{(k_1, k_2) : g(k_1, k_2) = \ell \text{ and } \mathbb{P}(X_1 = k_1, X_2 = k_2) > 0\}$. Observe that $\{g(X_1, X_2) = \ell\} = \{(X_1, X_2) \in A_{\ell}\}$. Remember that the left hand side is by definition

$$\mathbb{E}\left[g(X_1, X_2)\right] = \sum_{\ell: |A_{\ell}| > 0} \ell \cdot \mathbb{P}(g(X_1, X_2) = \ell) = \sum_{\ell: |A_{\ell}| > 0} \ell \cdot \mathbb{P}((X_1, X_2) \in A_{\ell}).$$

Finally, write $\ell \cdot \mathbb{P}((X_1, X_2) \in A_\ell) = \sum_{(k_1, k_2) \in A_\ell} g(k_1, k_2) \mathbb{P}[X_1 = k_1, X_2 = k_2].$

(b) $\mathbb{E}[X_1 + X_2] = \mathbb{E} X_1 + \mathbb{E} X_2$.

Hint: Express $X_1 + X_2$ as $g(X_1, X_2)$ for an appropriate choice of g and apply part (a).

(c) Show that if X_1 and X_2 are independent then

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[X_1] \mathbb{E}[X_2].$$

Hint: Express $\mathbb{E}[XY]$ as $\mathbb{E}[g(X,Y)]$ for an appropriate choice of $g:\mathbb{R}^2\to\mathbb{R}$. Use part (a) to conclude.

$$\sum_{k_1,k_2} (k_1 + k_2) p_{X_1,X_2}(k_1,k_2) = \sum_{k_1} \sum_{k_2} k_1 p_{X_1,X_2}(k_1,k_2) + \sum_{k_2} \sum_{k_1} k_2 p_{X_1,X_2}(k_1,k_2),$$

and $\sum_{k_1,k_2} a_{k_1} b_{k_2} = \sum_{k_1} a_{k_1} \sum_{k_2} b_{k_2}$.

¹Recall that a sum with non-negative summands can be split into two sums and can also be rearranged in an arbitrary fashion. The same is true for an absolutely convergent sum. E.g.,

- 4. Compute the moment generating functions of
 - (i) $X \sim \text{Geom}(p)$.

$$M(t) = \sum_{k=1}^{\inf} e^{kt} p (1-p)^{k-1}$$

$$= p(1-p)^{-1} \sum_{k=1}^{\inf} ((1-p)e^t)^k$$

$$= p(1-p)^{-1} \times \frac{(1-p)e^t}{1-qe^t}$$

$$= \frac{p(1-p)e^t}{(1-p)(1-(1-p)e^t)}$$

$$= \frac{pe^t}{1-(1-p)e^t}$$
Bounds: $(1-p)e^t < 1$

$$e^t < \frac{1}{1-p}$$

$$t < ln(\frac{1}{1-p})$$

$$t < -ln(1-p)$$

(ii) (Not marked) $Y \sim \text{Exp}(\lambda)$ and of $Z \sim \text{Poisson}(\mu)$.

5. (Only parts (c) and (d) will be marked)

(a) Show that if X and Y are independent random variables, and $f, g : \mathbb{R} \to \mathbb{R}$ then f(X) and g(Y) are also independent. Hint: By the definition of independence, it suffices to prove that for all $B, B' \subset \mathbb{R}$ we have that

$$\mathbb{P}[f(X) \in B, g(Y) \in B'] = \mathbb{P}[f(X) \in B] \mathbb{P}[g(Y) \in B'].$$

Note that $\{f(X) \in B\} = \{X \in \{x : f(x) \in B\}\}\$ and similarly, $\{g(Y) \in B\} = \{Y \in \{y : g(y) \in B'\}\}\$.

(b) Show that if X and Y are independent then the moment generating functions of X, Y and X + Y satisfy the following relation for all t:

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

Hint: Express $M_{X+Y}(t)$ as $\mathbb{E}[f(X)f(Y)]$ for an appropriate $f: \mathbb{R} \to \mathbb{R}_+$ and use part (a) of this question together with part (c) of Question 3 with g(x,y) = f(x)f(y).

(c) Let X and Y be two independent Poisson r.v.'s with parameters μ , λ . Compute the m.g.f. of X+Y and compare it with that of $W \sim \text{Poisson}(\mu+\lambda)$. Conclude that $X+Y \sim \text{Poisson}(\mu+\lambda)$.

$$\mathbb{P}(x) = e^{-\mu} \frac{\mu}{k!}
M_x(t) = \sum_{k \ge 0} e^{tk} e^{-\mu} \frac{\mu}{k!}
= e^{-\mu} e^{e^t} \mu
= e^{-\mu(e^t - 1)}
M_y(t) = \sum_{k \ge 0} e^{tk} e^{-\lambda} \frac{\lambda}{k!}
= e^{-\lambda} e^{e^t} \lambda
= e^{-\lambda(e^t - 1)}
M_{x+y}(t) = M_x(t) M_y(t)
= e^{-\mu(e^t - 1)} e^{-\lambda(e^t - 1)}
= e^{(\mu + \lambda)(e^t - 1)}
= M_w(t)$$

(d) Let $X \sim \mathcal{N}(a, b)$ and $Y \sim \mathcal{N}(c, d)$ be independent. Compute the m.g.f. of X + Y and compare it with that of $W \sim \mathcal{N}(a+c, b+d)$. Conclude that $X + Y \sim \mathcal{N}(a+c, b+d)$. Let $Z \sim \mathcal{N}(0, 1)$

$$f_z(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

$$M_z(t) = \int_{-\inf}^{\inf} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \times e^{tx} dx$$

$$= \int_{-\inf}^{\inf} \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2} + \frac{2tx}{2} + \frac{t^2}{2} - \frac{t^2}{2}} dx$$

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

$$= e^{\frac{(t)^2}{2}} \times 1 \to p.d.f. \text{ of } N(t, 1)$$

$$= e^{\frac{(t)^2}{2}}$$

Let $V \sim N(\mu, \sigma^2)$

$$M_v(t) = \mathbb{E}[e^{t(\sigma T + \mu)}]$$

$$= e^{t\mu} \mathbb{E}[e^{t(\sigma T)}]$$

$$= e^{t\mu} M_z(t\sigma)$$

$$= e^{t\mu} e^{\frac{(t\sigma)^2}{2}}$$

$$M_{X+Y}(t) = M_x(t)M_y(t)$$

$$= e^{ta + \frac{(tb)^2}{2}} e^{tc + \frac{(td)^2}{2}}$$

$$= e^{ta + tc + \frac{t^2(b^2 + d^2)}{2}}$$

$$= e^{t(a+c) + \frac{t^2(b^2 + d^2)}{2}}$$

$$= M_w(t)$$

Remarks: We will later be able to give alternative proofs for the conclusions of parts (c) and (d) using the convolution formula.

6. Suppose that $X \sim \text{Exp}(\mu)$. Find the probability density function of $Y = \ln(X)$.

$$\mathbb{P}(X \le x) = 1 - e^{-\lambda x}$$

$$\mathbb{P}(X \le e^y) = 1 - e^{-\lambda e^y}$$

$$F_y(Y) = \mathbb{P}(Y \le y) = \mathbb{P}(\ln(X) \le y) = \mathbb{P}(X \le e^y)$$

$$= 1 - e^{-\lambda e^y}$$

$$f_y(Y) = \lambda e^y e^{-\lambda e^y} = \lambda e^{y - \lambda e^y}$$

7. Let $X \sim \text{Bin}(n, p)$. Find the moment generating function $M_X(t)$ of X. Use your answer to calculate $\mathbb{E}[X]$ and Var(X).

Hint: Consider $Y = \sum_{i=1}^n \xi_i$ where ξ_1, \ldots, ξ_n are i.i.d. Bernoulli(p) random variables. Then X and Y have the same distribution and so $M_X = M_Y$, $\mathbb{E}[X^n] = \mathbb{E}[Y^n]$ for all n and $\mathrm{Var}(X) = \mathrm{Var}(Y)$. You may use the following generalization of 5(b): $M_Y(t) = M_{\sum_{i=1}^n \xi_i}(t) = \prod_{j=1}^n M_{\xi_j}(t)$ (by independence of ξ_1, \ldots, ξ_n).

$$M(t) = \sum_{k} e^{kt} \binom{n}{k} p^{k} (1-p)^{n-k}$$

$$= \sum_{k} \binom{n}{k} (pe^{t})^{k} (1-p)^{n-k}$$

$$= (pe^{t} + (1-p))^{n}$$

$$M'(t) = n(pe^{t} + (1-p))^{n-1} (pe^{t})$$

$$M''(t) = n(n-1)(pe^{t} + (1-p))^{n-2} (pe^{t})^{2} + n(pe^{t} + (1-p))^{n-1} (pe^{t})$$

$$\mathbb{E}[X] = M'(0) = np$$

$$Var(X) = M''(0) - (M'(0))^{2}$$

$$= n(n-1)(p-p)^{n-2}p^{2} + n(p-p)^{n-1}p - (np)^{2}$$

$$= np(1 + (n-1)p) - (np)^{2}$$

$$= np(1 + (n-1)p - np)$$

$$= np(1 + np - p - np)$$

$$= np(1-p)$$

8. (Not marked) Let X be a continuous random variable. Denote its PDF by f. Suppose that (i) f(x) = 0 for all x < 0, (ii) f(x) > 0 for all $x \ge 0$, (iii) f is right-continuous at 0 and continuous elsewhere, and (iv) for all $s, t \ge 0$

$$\mathbb{P}[X>t+s\mid X>s]=\mathbb{P}[X>t]\quad \text{(memoryless property for a continuous r.v.)}.$$

Show that $X \sim \text{Exp}(\lambda)$ where $\lambda := f(0)$.

Hint: Use the memoryless property in order to derive a differential equation for $G = 1 - F_X$ and solve it.

Remark: We proved in lecture that if $X \sim \text{Exp}(\lambda)$ then the above holds. This exercise shows that among random variables satisfying (i)-(iii), the memoryless property characterizes the Exponential family of distributions.

9. Challenge (not marked): Let X and Y be continuous random variables defined on the same probability space. Denote their joint cumulative distribution function by

$$F_{X,Y}(x,y) := \mathbb{P}[X \le x, Y \le y].$$

Let $S \subseteq \mathbb{R}^2$ be an open set such that the partial derivatives $\partial_1 F_{X,Y}$, $\partial_2 F_{X,Y}$, $\partial_1^2 F_{X,Y}$, $\partial_1 \partial_2 F_{X,Y}$ and $\partial_2 \partial_1 F_{X,Y}$ exist and are continuous on S. Show that for all $(s,t) \in S$ and all $\varepsilon, \delta > 0$ such that $B := (s, s + \varepsilon] \times (t, t + \delta] \subseteq S$ we have that

$$\mathbb{P}[(X,Y) \in B] = \int_{B} \partial_1 \partial_2 F_{X,Y}(x,y) dx dy.$$

Remark: It follows that if $\mathbb{P}[(X,Y) \in S] = 1$, then X,Y have a joint density function

$$f_{X,Y}(x,y) := \begin{cases} \partial_1 \partial_2 F_{X,Y}(x,y) = \partial_2 \partial_1 F_{X,Y}(x,y) & (x,y) \in S \\ 0 & (x,y) \notin S \end{cases}.$$

Hint: Use the relation from lecture: For all $s, t \in \mathbb{R}$ and all $\varepsilon, \delta \geq 0$ we have that

$$\mathbb{P}[X \in (s, s + \varepsilon], Y \in (t, t + \delta]] = F_{X,Y}(s + \varepsilon, t + \delta) + F_{X,Y}(s, t) - F_{X,Y}(s + \varepsilon, t) - F_{X,Y}(s, t + \delta).$$

10. Challenge (not marked): Let X_1, \ldots, X_n be continuous random variables defined on the same probability space. Denote the probability density function of X_i by f_i . Assume that for all closed sets $B \subset \mathbb{R}^n$ we have that

$$\mathbb{P}[(X_1,\ldots,X_n)\in B]=\int_B f_1(x_1)f_2(x_2)\ldots f_n(x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n.$$

Show that X_1, \ldots, X_n are independent.

Hint: By a fact stated in lecture, X_1, \ldots, X_n are independent if and only if for all $t_1, \ldots, t_n \in \mathbb{R}$ we have that $\{X_1 \leq t_1\}, \{X_2 \leq t_2\}, \ldots$ and $\{X_n \leq t_n\}$ are independent.

11. Challenge (not marked): Let X_1, \ldots, X_n be continuous random variables defined on the same probability space.

(a) Show that for all $i \in [n] := \{1, 2, ..., n\}$ and $s \in \mathbb{R}$

$$\lim_{t\to\infty} \mathbb{P}[\{X_i\leq s\}\cap\{X_j\leq t \text{ for all } j\in[n]\setminus\{i\}\}] = \mathbb{P}[X_i\leq s] \quad \text{and that}$$

$$\lim_{t \to -\infty} \mathbb{P}[\{X_i > s\} \cap \{X_j > t \text{ for all } j \in [n] \setminus \{i\}\}] = \mathbb{P}[X_i > s].$$

Hint: Use the hint from challenge question 12 in HW assignment 5. Namely, that for every nested sequence of events $B_1 \subset B_2 \subset \cdots$ we have that $\mathbb{P}(\bigcup_{m=1}^{\infty} B_m) = \lim_{m \to \infty} \mathbb{P}(B_m)$.

(b) Assume that there are some $G_1, \ldots, G_n : \mathbb{R} \to \mathbb{R}_+$ such that: (i) $\lim_{t \to -\infty} G_j(t) = 1$ for all $j \in [n]$, and (ii) for all $t_1, \ldots, t_n \in \mathbb{R}$ we have that

$$\mathbb{P}[X_1 > t_1, X_2 > t_2 \dots, X_n > t_n] = G_1(t_1)G_2(t_2)\cdots G_n(t_n).$$

Show that X_1, \ldots, X_n are independent and that $\mathbb{P}[X_i > t] = G_i(t)$ for all $1 \le i \le n$ and $t \in \mathbb{R}$.

Hint: By a fact stated in lecture, X_1, \ldots, X_n are independent if and only if for all $t_1, \ldots, t_n \in \mathbb{R}$ we have that $\{X_1 \leq t_1\}, \{X_2 \leq t_2\}, \ldots$ and $\{X_n \leq t_n\}$ are independent. By another fact from lecture, for every n independent events A_1, \ldots, A_n and every $i \in [n]$ we have that $A_1^c, \ldots, A_i^c, A_{i+1}, \ldots, A_n$ are independent. Apply this with i = n for some appropriate events.

12. Challenge (not marked): Let X and Y be two independent random variables. Denote their CDFs by F_X and F_Y , respectively. Assume that F_X and F_Y are continuous functions. Show that

$$\mathbb{P}[X=Y]=0.$$

13. Let X_1, X_2 be independent $\text{Exp}(\lambda)$ random variables. Let $W_1 := \min\{X_1, X_2\}$ and $W_2 := \max\{X_1, X_2\} - \min\{X_1, X_2\}$. Show that $W_1 \sim \text{Exp}(2\lambda), W_2 \sim \text{Exp}(\lambda)$ and that W_1 and W_2 are independent.

Hint: You may use the following without further justification, beyond what is stated below. By Question 11 (b), it suffices to show that

$$\forall s, t \ge 0, \quad \mathbb{P}[W_1 > s, W_2 > t] = \mathbb{P}[\text{Exp}(2\lambda) > s]\mathbb{P}[\text{Exp}(\lambda) > t] = e^{-2\lambda s}e^{-\lambda t}.$$

By Question 12 $\mathbb{P}[W_2 > 0] = \mathbb{P}[X_1 \neq X_2] = 1$. This can be used to verify the last display when t = 0 (using a theorem from lecture concerning the distribution of W_1), and that for t > 0 this implies that

$$\begin{split} \mathbb{P}[W_1 > s, W_2 > t] &= \mathbb{P}[W_1 > s, W_2 > t, X_1 < X_2] + \mathbb{P}[W_1 > s, W_2 > t, X_2 < X_1] \\ &= \mathbb{P}[X_1 > s, X_2 - X_1 > t, X_1 < X_2] + \mathbb{P}[X_2 > s, X_1 - X_2 > t, X_2 < X_1] \\ &= \mathbb{P}[X_1 > s, X_2 > t + X_1] + \mathbb{P}[X_2 > s, X_1 > t + X_2] \end{split}$$

By symmetry, the two probabilities are equal to one another. Consider

 $X_1, X_2 > 0 \Longrightarrow W_1 > 0$

$$B := \{(x, y) \in \mathbb{R}^2 : x > s \quad \text{and} \quad x + t < y < \infty\},$$

and observe that $\{X_1 > s, X_2 > t + X_1\} = \{(X_1, X_2) \in B\}$. Hence, using the definition of a joint probability density function and Fubini's theorem

$$\mathbb{P}[X_1 > s, X_2 > t + X_1] = \mathbb{P}[(X_1, X_2) \in B] = \iint_B f_{X_1, X_2}(x, y) dx dy = \int_s^{\infty} \int_{t+x}^{\infty} f_{X_1, X_2}(x, y) dy dx,$$

where f_{X_1,X_2} is the joint probability density function of X_1, X_2 . Finally, by independence and a fact from Section 6.3 $f_{X_1,X_2}(x,y) = f_{X_1}(x)f_{X_2}(y) = \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda y}$.

$$\begin{split} \mathbb{P}(W_{1} > s) &= \mathbb{P}(\min(X_{1}, X_{2}) > s) \\ &= \mathbb{P}(X_{1} > s) \mathbb{P}(X_{2} > s) \\ &= e^{-\lambda s} e^{-\lambda s} \\ &= e^{-2\lambda s} \\ F_{w_{1}}(s) &= \mathbb{P}(W_{1} \le s) = \mathbb{P}(W_{1} > s) \\ &= 1 - e^{-2\lambda s} \\ f_{w_{1}}(s) &= 2\lambda e^{-2\lambda s} \\ &\sim Exp(2\lambda) \end{split}$$

$$f_{W_{2}} &= \int_{0}^{\infty} f_{X_{1}}(X_{2} + W_{2}) f_{X_{2}}(X_{2}) dX_{2} + \int_{t}^{\infty} f_{X_{1}}(X_{2} - t) f_{X_{2}}(X_{2}) dX_{2} \\ &= \int_{0}^{\infty} \lambda e^{-\lambda X_{2}} dX_{2} + \int_{t}^{\infty} \lambda e^{-\lambda X_{2}} dX_{2} \\ &= [-e^{-\lambda y}]_{0}^{\infty} [-e^{-\lambda y}]_{t}^{\infty} \\ &= e^{-\lambda t} \\ &\sim Exp(\lambda) \end{split}$$

Showing Independence,

$$\mathbb{P}[W_1 > s, W_2 > t] = \mathbb{P}[\text{Exp}(2\lambda) > s] \mathbb{P}[\text{Exp}(\lambda) > t]$$
$$= \mathbb{P}[W_1 > s] \mathbb{P}[W_2 > t]$$
$$= e^{-2\lambda s} e^{-\lambda t}$$