STAT3035/8035
Tutorial 1

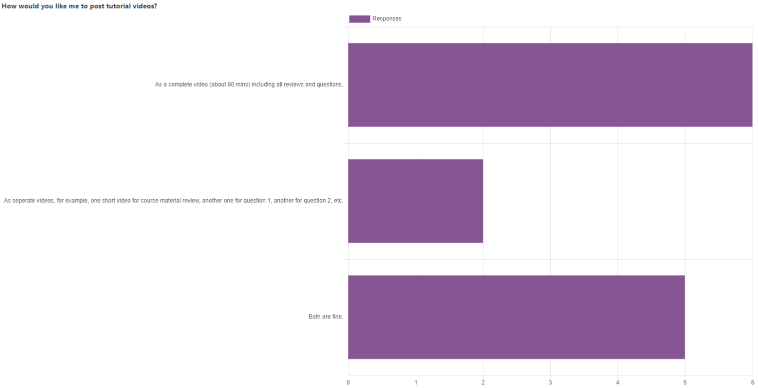
Marco Li

Contact: qingyue.li@anu.edu.au

Outline

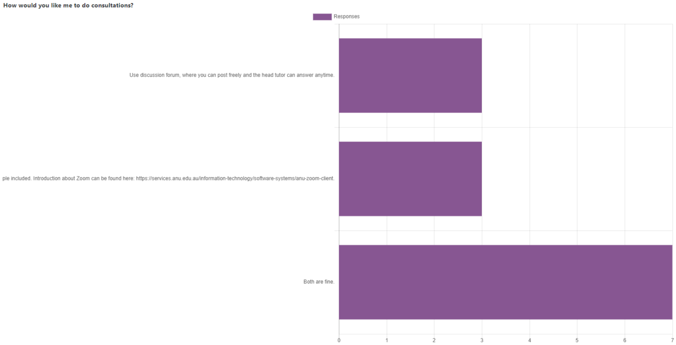
Tutorial arrangement

- Tutorial update time: Monday each week
- Tutorial form: one video including everything



Consultation arrangement

- Consultation form: Discussion forum or Zoom by appointment



- Post your questions in corresponding discussion post, or
- I will reserve 5:30pm - 6:00pm (Canberra time) each Friday as a Zoom consultation time, **IF** there are student(s) making appointment by the end of each Thursday in discussion forum or by email

Some useful suggestions

- Watch the lecture recordings right after they come out - don't procrastinate
- Read the lecture slides
- Do tutorial questions before watching my tutorial video

Tutorial plan

- Review related lecture materials
- Tutorial questions

Outline

Probability theory

- pdf/pmf of known distributions
  - discrete: Binomial, Poisson...
  - continuous: Gamma, Normal...
- Moments (raw) - EX^n
- Quantiles - e.g. median

- MOM
- Solve system:

$$E_\theta(X) = \bar{x} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i$$
$$\vdots$$
$$E_\theta\left(X^k\right) = \overline{x^k} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i^k$$

where k is number of parameters.

- MOP
- Solve system:

$$x_{p_1} = \hat{x}_{p_1}$$
$$\vdots$$
$$x_{p_k} = \hat{x}_{p_k}$$

where $\hat{x}_p$ is observed p^{th} percentile of data, for some choice of $p_1, \ldots, p_k$

Statistics - Estimation

- MLE
- Likelihood Function: $L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f_X(x_i; \theta)$
- Log-Likelihood Function: $l(\theta) = \ln\{L(\theta; x_1, \ldots, x_n)\}$
- Maximum Likelihood Estimate, $\hat{\theta}_{MLE}$, solves:

$$\frac{\partial l(\theta)}{\partial \theta_i} = 0, \quad 1 \leq i \leq k$$

- Maximum Likelihood Theorem: For large samples,

$$\Pr_\theta \left\{ \frac{\hat{\theta}_{MLE} - \theta}{\sqrt{I^{-1}\left(\hat{\theta}_{MLE}\right)}} \leq t \right\} \approx \Phi(t)$$

  where $\Phi(\cdot)$ is the standard normal CDF and $I(\theta) = -E_\theta\{l''(\theta)\}$

- So, we can use $\hat{\theta}_{MLE} \pm 1.96\sqrt{I^{-1}\left(\hat{\theta}_{MLE}\right)}$ as an approximate 95% confidence interval for θ

Statistics - Goodness of fit testing

- Pearson Chi-Squared Test
- Idea:
  - Data is n iid observations classified into k categories
  - O_i = number of observations in category i
  - "Theory": $p_i = \Pr(\text{ obs. in cat. } i)$
  - $E_i = np_i$ = expected # of obs. in ith category
  - Measure discrepancy using test statistic:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

    which has an approximate χ^2 -distribution with a number of degrees of freedom equal to:
    $df = k - 1 - (\#\text{ parameters estimated in determining } p_i)$
- Implementation:
  - For continuous data, need to "discretise" using bins
  - Choose 5 to 15 bins
  - Could use histogram bins (equal width)

Outline

Question 1

(Joint Distributions and Independency) Let X and Y be discrete random variables each taking values in the sample space $S = \{0, 1, 2\}$ and having a joint probability mass funct ion given by the following table:

		Y		
		0	1	2
X	0	0.10	0.10	0.20
	1	0.25	0.00	0.20
	2	0.05	0.05	0.05

(a)* Find the probability mass function and cumulative distribution function of $U := X + Y$.
(b) Find the marginal probability mass functions, a.k.a unconditional probability mass functions, of both X and Y. Are X and Y independent?

Solution 1

Solution 1

(a) Clearly the sample space for U is $S_U = \{0,1,2,3,4\}$, and we can easily calculate:
$p_U(0) = \mathbb{P}(U=0) = \mathbb{P}\{(X=0, Y=0)\} = 0.1$
$p_U(1) = \mathbb{P}(U=1) = \mathbb{P}\{(X=0, Y=1) \text{ or } (X=1, Y=0)\} = 0.1 + 0.25 = 0.35$
$p_U(2) = \mathbb{P}(U=2) = \mathbb{P}\{(X=0, Y=2) \text{ or } (X=1, Y=1) \text{ or } (X=2, Y=0)\} =$
$0.2 + 0 + 0.05 = 0.25$
$p_U(3) = \mathbb{P}(U=3) = \mathbb{P}\{(X=1, Y=2) \text{ or } (X=2, Y=1)\} = 0.2 + 0.05 = 0.25$
$p_U(4) = \mathbb{P}(U=4) = \mathbb{P}\{(X=2, Y=2)\} = 0.05$

(b) The marginal probability mass function of X is:
$p_X(0) = \mathbb{P}(X=0) = \mathbb{P}(X=0, 0 \le Y \le 2) = 0.10 + 0.10 + 0.20 = 0.40$
$p_X(1) = \mathbb{P}(X=1) = \mathbb{P}(X=1, 0 \le Y \le 2) = 0.25 + 0.00 + 0.20 = 0.45$
$p_X(2) = \mathbb{P}(X=2) = \mathbb{P}(X=2, 0 \le Y \le 2) = 0.05 + 0.05 + 0.05 = 0.15$
And similarly the probability mass function of Y is:
$p_Y(0) = \mathbb{P}(Y=0) = \mathbb{P}(Y=0, 0 \le X \le 2) = 0.10 + 0.25 + 0.05 = 0.40$
$p_Y(1) = \mathbb{P}(Y=1) = \mathbb{P}(Y=1, 0 \le X \le 2) = 0.10 + 0.00 + 0.05 = 0.15$
$p_Y(2) = \mathbb{P}(Y=2) = \mathbb{P}(Y=2, 0 \le X \le 2) = 0.20 + 0.20 + 0.05 = 0.45$
Now, clearly X and Y are not independent since, for example,
$\mathbb{P}(X=0, Y=0) = 0.1 \neq 0.16 = \mathbb{P}(X=0)\mathbb{P}(Y=0)$

Question 1

(Joint Distributions and Independency) Let X and Y be discrete random variables each taking values in the sample space $S = \{0, 1, 2\}$ and having a joint probability mass funct ion given by the following table:

		Y		
		0	1	2
X	0	0.10	0.10	0.20
	1	0.25	0.00	0.20
	2	0.05	0.05	0.05

(c)* Let X_1 be a discrete random variable having a probability mass function equal to the marginal probability mass function of X calculated in part (b). Similarly, let Y_1 be a discrete random variable having a probability mass function equal to the marginal probability mass function of Y calculated in part (b). Also, let X_1 and Y_1 be independent. Construct a table similar to the one above giving the joint probability mass function of X_1 and Y_1.
(d)* Using your result from part (c), calculate the probability mass function of the random variable $U_1 := X_1 + Y_1$. Compare this probability mass function with the one you calculated in part (a).
(e)* Compute $\mathbb{E}X$ and $\mathbb{V}Y$.

Solution 1

Solution 1

(c) Solution: Using the multiplication rule for independent random variables, we have:

		Y		
		0	1	2
X	0	$0.40 \times 0.40 = 0.16$	$0.40 \times 0.15 = 0.0600$	$0.40 \times 0.45 = 0.1800$
	1	$0.45 \times 0.40 = 0.18$	$0.45 \times 0.15 = 0.0675$	$0.45 \times 0.45 = 0.2025$
	2	$0.15 \times 0.40 = 0.06$	$0.15 \times 0.15 = 0.0225$	$0.15 \times 0.45 = 0.0675$

(d) Solution: Similar to part (a), the probability mass function of U_1 is calculated as:

$$
\begin{aligned}
p_{U_1}(0) &= \mathbb{P}(U_1 = 0) = \mathbb{P}\{(X_1 = 0, Y_1 = 0)\} = 0.16; \\
p_{U_1}(1) &= \mathbb{P}(U_1 = 1) = \mathbb{P}\{(X_1 = 0, Y_1 = 1) \text{ or } (X_1 = 1, Y_1 = 0)\} = 0.06 + 0.18 = 0.24 \\
p_{U_1}(2) &= \mathbb{P}(U_1 = 2) = \mathbb{P}\{(X_1 = 0, Y_1 = 2) \text{ or } (X_1 = 1, Y_1 = 1) \text{ or } (X_1 = 2, Y_1 = 0)\} \\
&= 0.18 + 0.0675 + 0.06 = 0.3075 \\
p_{U_1}(3) &= \mathbb{P}(U_1 = 3) = \mathbb{P}\{(X_1 = 1, Y_1 = 2) \text{ or } (X_1 = 2, Y_1 = 1)\} = 0.2025 + 0.0225 = 0.225 \\
p_{U_1}(4) &= \mathbb{P}(U_1 = 4) = \mathbb{P}\{(X_1 = 2, Y_1 = 2)\} = 0.0675
\end{aligned}
$$

This is different from the probability mass function of U, despite the equality of the marginal distributions of the components of U_1 and U. Thus, the joint distribution is necessary in determining the distribution of the sum (or any multi-variable function) of random variables.

Solution 1

(e) By definition we calculate:

$$\mathbb{E}X = \sum_{i=0}^{2} i p_X(i) = 0 \times 0.4 + 1 \times 0.45 + 2 \times 0.15 = 0.75$$

$$\mathbb{E}Y = \sum_{i=0}^{2} i p_Y(i) = 0 \times 0.4 + 1 \times 0.15 + 2 \times 0.45 = 1.05$$

$$\mathbb{E}Y^2 = \sum_{i=0}^{2} i^2 p_Y(i) = 0^2 \times 0.4 + 1^2 \times 0.15 + 2^2 \times 0.45 = 1.95$$

$$\mathbb{V}Y = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = 1.95 - 1.05^2 = 0.8475.$$

Question 1

(Joint Distributions and Independency) Let X and Y be discrete random variables each taking values in the sample space $S = \{0, 1, 2\}$ and having a joint probability mass funct ion given by the following table:

		Y		
		0	1	2
X	0	0.10	0.10	0.20
	1	0.25	0.00	0.20
	2	0.05	0.05	0.05

(f) Calculate the probability mass function of the random variable $\mathbb{E}(X|Y)$ by finding $\mathbb{E}(X|Y = y)$ and calculate $\mathbb{P}[\mathbb{E}(X|Y) = \mathbb{E}(X|Y = y)]$ for each of $y = 0, 1, 2$. Verify the identity $\mathbb{E}(X) = \mathbb{E}[\mathbb{E}(X|Y)]$ for these two random variables.
(g) In addition to (f), verify the identity $\mathbb{V}X = \mathbb{E}[\mathbb{V}(X|Y)] + \mathbb{V}[\mathbb{E}(X|Y)]$ for these two random variables.

Solution 1

Solution 1

(f) We note that $\mathbb{E}(X|Y=y)=\sum_{i=0}^{2} i p_{X|Y}(i|y)$ where $p_{X|Y}(i|y)$ is the conditional probability mass function of X given $Y=y$ which we can calculate for all possible pairs (x,y) as:

$p_{X|Y}(0|0)=\mathbb{P}(X=0|Y=0)=\frac{\mathbb{P}(X=0,Y=0)}{\mathbb{P}(Y=0)}=\frac{0.10}{0.40}=0.250$

$p_{X|Y}(1|0)=\mathbb{P}(X=1|Y=0)=\frac{\mathbb{P}(X=1,Y=0)}{\mathbb{P}(Y=0)}=\frac{0.25}{0.40}=0.625$

$p_{X|Y}(2|0)=\mathbb{P}(X=2|Y=0)=\frac{\mathbb{P}(X=2,Y=0)}{\mathbb{P}(Y=0)}=\frac{0.05}{0.40}=0.125$

$p_{X|Y}(0|1)=\mathbb{P}(X=0|Y=1)=\frac{\mathbb{P}(X=0,Y=1)}{\mathbb{P}(Y=1)}=\frac{0.10}{0.15}=0.667$

$p_{X|Y}(1|1)=\mathbb{P}(X=1|Y=1)=\frac{\mathbb{P}(X=1,Y=1)}{\mathbb{P}(Y=1)}=\frac{0}{0.15}=0.000$

$p_{X|Y}(2|1)=\mathbb{P}(X=2|Y=1)=\frac{\mathbb{P}(X=2,Y=1)}{\mathbb{P}(Y=1)}=\frac{0.05}{0.15}=0.333$

$p_{X|Y}(0|2)=\mathbb{P}(X=0|Y=2)=\frac{\mathbb{P}(X=0,Y=2)}{\mathbb{P}(Y=2)}=\frac{0.20}{0.45}=0.444$

$p_{X|Y}(1|2)=\mathbb{P}(X=1|Y=2)=\frac{\mathbb{P}(X=1,Y=2)}{\mathbb{P}(Y=2)}=\frac{0.20}{0.45}=0.444$

$p_{X|Y}(2|2)=\mathbb{P}(X=2|Y=2)=\frac{\mathbb{P}(X=2,Y=2)}{\mathbb{P}(Y=2)}=\frac{0.05}{0.45}=0.111$

Thus, we can calculate

$$
\begin{aligned}
\mathbb{E}(X|Y=0)&=0\times 0.250+1\times 0.625+2\times 0.125=0.875\\
\mathbb{E}(X|Y=1)&=0\times 0.667+1\times 0.000+2\times 0.333=0.667\\
\mathbb{E}(X|Y=2)&=0\times 0.444+1\times 0.444+2\times 0.111=0.667
\end{aligned}
$$

Finally, then, we see that
$\mathbb{E}[\mathbb{E}(X|Y)]=\sum_{i=0}^{2}\mathbb{E}(X|Y=i)p_Y(i)=0.875\times 0.4+0.667\times 0.15+0.667\times 0.45=0.75$
which is the same as $\mathbb{E}(X)$ which we calculated in part (e).

Solution 1

(g) First by (e) and definition we have

$$\mathbb{E}X = 0.75$$

$$\mathbb{E}X^2 = \sum_{i=0}^{2} i^2 p_X(i) = 0^2 \times 0.15 + 1^2 \times 0.45 + 2^2 \times 0.45 = 1.05$$

$$\mathbb{V}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 1.05 - 0.75^2 = 0.4875$$

Now by (f) and definition we have
$\mathbb{E}\left(X^2|Y=0\right) = 0^2 \times 0.250 + 1^2 \times 0.625 + 2^2 \times 0.125 = 1.125$
$\mathbb{E}\left(X^2|Y=1\right) = 0^2 \times 0.667 + 1^2 \times 0.000 + 2^2 \times 0.333 = 1.333$
$\mathbb{E}\left(X^2|Y=2\right) = 0^2 \times 0.444 + 1^2 \times 0.444 + 2^2 \times 0.111 = 0.888$
$\mathbb{V}(X|Y=0) = \mathbb{E}\left(X^2|Y=0\right) - [\mathbb{E}(X|Y=0)]^2 = 1.125 - 0.875^2 = 0.359$
$\mathbb{V}(X|Y=1) = \mathbb{E}\left(X^2|Y=1\right) - [\mathbb{E}(X|Y=1)]^2 = 1.333 - 0.667^2 = 0.889$
$\mathbb{V}(X|Y=2) = \mathbb{E}\left(X^2|Y=2\right) - [\mathbb{E}(X|Y=2)]^2 = 0.888 - 0.667^2 = 0.444$
$\mathbb{E}[\mathbb{V}(X|Y)] = \sum_{i=0}^{2} \mathbb{V}(X|Y=i)p_Y(i) = 0.359 \times 0.4 + 0.889 \times 0.15 + 0.444 \times 0.45 = 0.477$
$\mathbb{E}[\mathbb{E}(X|Y)]^2 = \sum_{i=0}^{2}[\mathbb{E}(X|Y=i)]^2 p_Y(i) = 0.875^2 \times 0.4 + 0.667^2 \times 0.15 + 0.667^2 \times 0.45 =$
0.57292
$\mathbb{V}[\mathbb{E}(X|Y)] = \mathbb{E}[\mathbb{E}(X|Y)]^2 - \{\mathbb{E}[\mathbb{E}(X|Y)]\}^2 = 0.57292 - 0.75^2 = 0.01042$
$\mathbb{V}X = 0.477 + 0.01042 = 0.4875$

Question 2

Let X be normally distributed with mean μ and variance σ^2. Find the moment generating function of $X, m_X(t) = \mathbb{E}\left(e^{tX}\right)$.

Solution 2

Solution 2

By using the technique discussed in the above remark we have

$$
\begin{aligned}
m_X(t) := \mathbb{E}\left(e^{tX}\right) &= \int_{\mathbb{R}} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2-2\left(\mu+\sigma^2 t\right)x+\mu^2}{2\sigma^2} dx} dx \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left[x-\left(\mu+\sigma^2 t\right)\right]^2-2\mu\sigma^2 t-\sigma^4 t^2}{2\sigma^2} dx} \\
&= e^{\mu t+\frac{\sigma^2}{2}t^2} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left[x-\left(\mu+\sigma^2 t\right)\right]^2}{\sqrt{2\pi}\sigma}} dx \\
&= e^{\mu t+\frac{1}{2}\sigma^2 t^2}
\end{aligned}
$$

Question 3

(Indicator Functions and Reinsurance) Let X be an exponential random variable with mean λ^{-1}.
(a) It is straightforward to show that the moment generation function of X is $m_X(t) = \lambda(\lambda - t)^{-1}$ for $t < \lambda$. Use this fact to find the moment generating function of $Y := X - m$ for a fixed constant m.
(b) Define the new random variable

$$Z := \begin{cases} X & \text{if} \quad X \leq M \\ M & \text{if} \quad X > M \end{cases}$$

for a fixed constant M. Note that an equivalent definition of Z is $Z = XI_{(X \leq M)} + MI_{(X > M)}$, where the function $I_{(\cdot)}$ is called an indicator function and is defined to be 1 if its argument is true and 0 other wise. Find the moment generating function of Z.

Solution 3

Solution 3

(a) As an exercise, prove the moment generating function formula given in the question. With the moment generating function given in the question, one can simply get the following

result.

$$m_Y(t) = \mathbb{E}e^{tY} = \mathbb{E}e^{t(X-m)} = e^{-tm}\mathbb{E}e^{tX} = \frac{e^{-tm}\lambda}{\lambda-t}, t < \lambda$$

(b) Apart from the new notation of indicator functions, this question is simply beer and skittle.

$$\begin{aligned} m_Z(t) = \mathbb{E}e^{tZ} &= \mathbb{E}e^{t\left(XI_{\{X\leq M\}}+MI_{\{X>M\}}\right)} \\ &= \int_0^\infty e^{t\left(xI_{\{x\leq M\}}+MI_{\{x>M\}}\right)}\lambda e^{-\lambda x}dx \\ &= \int_0^M e^{tx}\lambda e^{-\lambda x}dx + \int_M^\infty e^{tM}\lambda e^{-\lambda x}dx \\ &= \frac{\lambda e^{(t-\lambda)M}}{t-\lambda} - \frac{\lambda}{t-\lambda} + e^{(t-\lambda)M} \\ &= \frac{te^{(t-\lambda)M}}{t-\lambda} - \frac{\lambda}{t-\lambda} \end{aligned}$$

Question 4

(Distribution of a function of random variables) Let X be a continuous random variable having the following density function:

$$f(x) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad x > 0$$

The distribution having this density is often referred to as the "folded normal". Let $Y = X^2$. Find the density function of Y. Do you recognise the density you found?

Solution 4

Solution 4

First find cumulative distribution function and then find the corresponding density function by taking derivative.
CDF : $F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(X^2 \leq y\right) = \mathbb{P}(X \leq \sqrt{y}) = F_X(\sqrt{y})$
PDF:

$$\begin{aligned} f_y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(\sqrt{y}) = f_X(\sqrt{y}) \frac{d}{dy}\sqrt{y} \\ &= \sqrt{\frac{2}{\pi}} e^{-\frac{y}{2}} \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} = \frac{1}{\Gamma\left(\frac{1}{2}\right)\sqrt{2y}} e^{-\frac{y}{2}} \end{aligned}$$

Hence, $Y \sim \text{Gamma}\left(k = \frac{1}{2}, \theta = 2\right) = \chi^2_{(1)}$

Question 5

(Moment Generating Functions and Independency)
(a) Suppose that $X_1, \ldots, X_k$ are independent random variables with Gamma distributions having shape parameters $\alpha_i (i = 1, \ldots, k)$, respectively, and common scale parameter θ. Define $\alpha = \sum_{i=1}^{k} \alpha_i$ Prove that $X = \sum_{i=1}^{k} X_i$ has a Gamma distribution with parameters α and θ.
(b)* Suppose that X_1 and X_2 are independent random variables with Gamma distributions having common shape parameter α and scale parameters θ_1 and θ_2, respectively. Assuming that $\theta_1 \neq \theta_2$ do you think that $X := X_1 + X_2$ has a Gamma distribution? Why or why not?

Solution 5

Solution 5

(a) We note that the moment generating function of X_i is given by $m_{X_i}(t) = (1-\theta t)^{-\alpha_i}$ for $t < \theta^{-1}$. This then implies that the moment generating function of X is:

$$m_X(t) = \mathbb{E}\left(e^{tX}\right) = \mathbb{E}\left\{\exp\left(t\sum_{i=1}^{n} X_i\right)\right\} = \mathbb{E}\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} E\left(e^{tX_i}\right) = \prod_{i=1}^{k} m_{X_i}(t)$$
$$= (1-\theta t)^{-\sum_{i=1}^{k}\alpha_i} = (1-\theta t)^{-\alpha}, \quad t < \theta^{-1}$$

where the fourth equality follows from the assumed independence of the X_i 's. In the current case, we can easily recognise the calculated moment generating function as that of a Gamma distribution. So, X must be Gamma distributed with parameters α and θ.
(b) Let's examine the moment generating function of X, which has the form:

$$m_X(t) = m_{X_1}(t)m_{X_2}(t) = \{(1-\theta_1 t)(1-\theta_2 t)\}^{-\alpha}, \quad t < \min\left(\theta_1^{-1}, \theta_2^{-1}\right)$$

This certainly does not appear to have the form of the moment generating function of a Gamma distribution in general.