MTH 102: Probability and Statistics

Quiz 5

15/05/2020

Sanjit K. Kaul

Question 1. With probability 0.4 student marks are distributed as a Gaussian distribution with mean 10 and variance 1. With probability 0.6, the marks are distributed as a Gaussian distribution with mean 5 and variance 1. Independently of the marks obtained, the time taken by a student to finish an exam is, with probability 0.5, distributed as a uniform random variable over the range $(0, 10)$, and otherwise is distributed as an exponential random variable with mean 5. Note that the PDF of an exponential random variable with mean $1/\mu$ is given by $\mu e^{-\mu x},\ x \geq 0$.

Answer the following questions.

1) Derive the variance of an exponential random variable with mean 5. In the provided text box, state in words (in plain English) how you derived the variance. Also, write down the variance you calculated. Provide detailed steps that led to the answer later in the upload.
2) Derive the average time it takes a student to finish an exam. Write down the average and state how you calculated it in words in the text box and provide your detailed steps in the upload.
3) Derive the variance of the time it takes a student to finish an exam. Write down the variance and state how you calculated it in words in the text box and provide your detailed steps in the upload.
4) Derive the average marks obtained by a student. Write down the average and state how you calculated it in words in the text box and provide your detailed steps in the upload.
5) Derive the variance of the marks obtained by a student. Write down the variance and state how you calculated it in words in the text box and provide your detailed steps in the upload.
6) Derive the correlation of the marks obtained by a student and the time taken to finish the exam. Write down the correlation and state how you calculated it in words in the text box and provide your detailed steps in the upload.

Question 2. Consider two students. Suppose the marks of one of the students are distributed as a Gaussian distribution with mean 10 and variance 1 and the time taken to finish the exam by the student is distributed as a uniform random variable over the range $(0, 10)$. Let the marks of the other student be distributed as a Gaussian distribution with mean 5 and variance 1 and let the time taken to finish the exam by this student be distributed as an exponential random variable with mean 5. Assume that all the random variables above are independent of each other. Note that the PDF of an exponential random variable with mean $1/\mu$ is given by $\mu e^{-\mu x},\ x \geq 0$.

Derive the probability of the event that the marks obtained by the first student are less than the marks obtained by the second student and that the time taken by the first student is greater than the time taken by the second student. You can leave your answer in terms of the CDF or the CCDF of the standard normal. Explain your steps in the provided text box and provide the corresponding detailed math in the upload.

Question 1. With probability 0.4 student marks are distributed as a Gaussian distribution with mean 10 and variance 1. With probability 0.6, the marks are distributed as a Gaussian distribution with mean 5 and variance 1. Independently of the marks obtained, the time taken by a student to finish an exam is, with probability 0.5, distributed as a uniform random variable over the range $(0, 10)$, and otherwise is distributed as an exponential random variable with mean 5. Note that the PDF of an exponential random variable with mean $1/\mu$ is given by $\mu e^{-\mu x},\ x \geq 0$.

Answer the following questions.

10 1) Derive the variance of an exponential random variable with mean 5. In the provided text box, state in words (in plain English) how you derived the variance. Also, write down the variance you calculated. Provide detailed steps that led to the answer later in the upload.

15 2) Derive the average time it takes a student to finish an exam. Write down the average and state how you calculated it in words in the text box and provide your detailed steps in the upload.

3) Derive the variance of the time it takes a student to finish an exam. Write down the variance and state how you calculated it in words in the text box and provide your detailed steps in the upload.

15 4) Derive the average marks obtained by a student. Write down the average and state how you calculated it in words in the text box and provide your detailed steps in the upload.

5) Derive the variance of the marks obtained by a student. Write down the variance and state how you calculated it in words in the text box and provide your detailed steps in the upload.

6) Derive the correlation of the marks obtained by a student and the time taken to finish the exam. Write down the correlation and state how you calculated it in words in the text box and provide your detailed steps in the upload.

10

With prob $0.4 : N(10,1)$
$0.6 : N(5,1)$] Student Marks

With prob 0.5: $Unif(0,10)$
: $Exp(1/5)$] Time taken.

(1) Mean is given as 5. That is $\mu = 1/5$.

$Var[X] = E[X^2] - (E[X])^2$

We need to calculate

$$E[X^2] = \int_0^\infty x^2 \mu e^{-\mu x} dx.$$ [Integration By parts]

$$Var[X] = \frac{1}{\mu^2} = \frac{1}{(1/5)^2} = 25$$

(2) Let T be the time taken by a student.

$$E[T] = E[T \mid \text{Time is Unif Dist}]\, P[\text{Time is Unif Dist}] + E[T \mid \text{Time is Exp Dist}]\, P[\text{Time is Exp Dist}]$$

$$= \left(\frac{10+0}{2}\right) 0.5 + (5)(0.5)$$

[Note that Expected value of Unif (0,10) is 5.

(3) ~~$Var[T] = Var[T \mid \text{Time is Unif}]\, P[\text{Time is Unif}] + Var[T \mid \text{Time is Exp}]\, P[\text{Time is Exp}]$~~

~~$= \frac{(10-0)^2}{12} 0.5 + \frac{1}{25} 0.5$~~

Got to be careful with the variance operator.

Start with:

$Var[T] = E[T^2] - (E[T])^2$; You have $E[T]$. So can calculate $(E[T])^2$

You need $E[T^2] = E[T^2 \mid Unif]\, P[Unif] + E[T^2 \mid Exp]\, P[Exp] = \left(\frac{100}{12} + (5)^2\right) 0.5 + \frac{2}{(1/5)^2} (0.5)$

Side Note: Why $Var[T] \neq Var[T \mid A] P[A] + Var(T \mid A^c) P[A^c]$

From first principles (that is $f_T(t)dt \approx P[t \leq T < t+dt]$), we know:

$f_T(t) = f_{T|Unif}(t)\, P[Unif] + f_{T|Exp}(t)\, P[Exp]$

$Var[T] = \int (t - \mu_T)^2 f_T(t)\, dt$

$= \int (t^2 + \mu_T^2 - 2t\mu_T) f_T(t)\, dt$

$= \int t^2 f_T(t)\, dt + \int \mu_T^2 f_T(t)\, dt - 2 \int t \mu_T f_T(t)\, dt$

$= \left[\int t^2 [f_{T|Unif}(t)]\, dt\right] P(Unif) + \left[\int t^2 f_{T|Exp}(t)\, dt\right] P[Exp] - \mu_T^2$

$= Var[T \mid Unif]\, P(Unif) + (\mu_{T|Unif})^2 P(Unif) + Var[T \mid Exp] P[Exp] + (\mu_{T|Exp})^2 P[Exp] - \mu_T^2$

$\neq Var[T \mid Unif]\, P(Unif) + Var[T \mid Exp]\, P[Exp]$ //

(4) Let M be the marks.

$$E[M] = E[M \mid N(10,1)]\, P[\text{Marks distributed as } N(10,1)] + E[M \mid N(5,1)]\, P[\text{Marks dist as } N(5,1)]$$

$$= (10)(0.4) + (5)(0.6)$$

(5) $Var[M] = ?$ $E[M^2] = (1 + (10)^2) 0.4 + (1 + 5^2) 0.6 = (100)(0.4) + 25(0.6) + 1$

$(E[M])^2 = (100)(0.4)^2 + (25)(0.6)^2 + 2(10)(0.4)(5)(0.6)$.

$\Rightarrow Var[M] = 7$

(6) $E[MT] = ?$

Note M and T are independent.

$\therefore E[MT] = E[M]\, E[T]$.

We have already calculated $E[M]$ & $E[T]$

Question 2. Consider two students. Suppose the marks of one of the students are distributed as a Gaussian distribution with mean 10 and variance 1 and the time taken to finish the exam by the student is distributed as a uniform random variable over the range $(0, 10)$. Let the marks of the other student be distributed as a Gaussian distribution with mean 5 and variance 1 and let the time taken to finish the exam by this student be distributed as an exponential random variable with mean 5. Assume that all the random variables above are independent of each other. Note that the PDF of an exponential random variable with mean $1/\mu$ is given by $\mu e^{-\mu x}$, $x \geq 0$.

Derive the probability of the event that the marks obtained by the first student are less than the marks obtained by the second student and that the time taken by the first student is greater than the time taken by the second student. You can leave your answer in terms of the CDF or the CCDF of the standard normal. Explain your steps in the provided text box and provide the corresponding detailed math in the upload.

$$\boxed{\begin{array}{ll} M_1 \sim N(10,1) & M_2 \sim N(5,1) \\ T_1 \sim Unif(0,10) & T_2 \sim Exp(1/5) \end{array}}$$

Clearly Defining the RV(s) : 5/50.

Event of Interest is $\{M_1 < M_2\} \cap \{T_1 > T_2\}$

[Identify the event 10/50]

$$P[\{M_1 < M_2\} \cap \{T_1 > T_2\}]$$
$$= P[M_1 < M_2]\ P[T_1 > T_2]$$

{ Applying Independence

10/50

$$P[M_1 < M_2]$$
$$= P[M_1 - M_2 < 0].$$

[12.5/50]

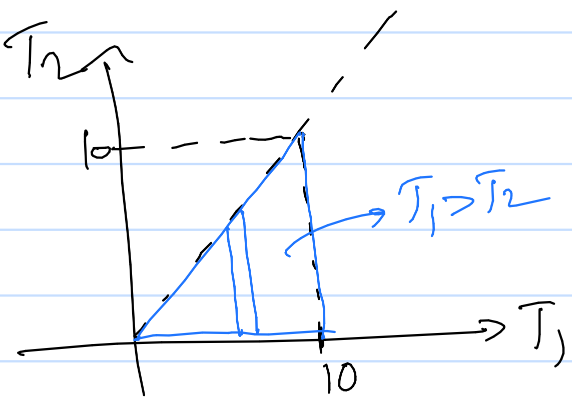
Note that $M_1 - M_2$ is Gaussian$(5,1)$

Let $Z = M_1 - M_2$. $Z \sim$ Gaussian$(5, 2)$; $Var[Z] = 2$

$$P[M_1 - M_2 < 0] = P[Z < 0]$$
$$= P\left[\frac{Z-5}{\sqrt{2}} < \frac{-5}{\sqrt{2}}\right]$$
$$= \Phi\left(\frac{-5}{\sqrt{2}}\right),$$ where $\Phi(x)$ is the CDF of the standard normal.

$$P(T_1 > T_2)$$ [12.5/50]

$$= \int_0^{10} \int_0^{t_1} f_{T_1,T_2}(t_1,t_2)\, dt_2\, dt_1$$



$$f_{T_1,T_2}(t_1,t_2) = f_{T_1}(t_1)\, f_{T_2}(t_2)$$ [Since T_1 & T_2 are given to be independent]

$$= \begin{cases} \frac{1}{5} e^{-t_2/5} \left(\frac{1}{10}\right) & 0 \leq t_1 \leq 10,\ 0 \leq t_2 \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore P(T_1 > T_2)$$
$$= \int_0^{10} \frac{1}{10} \int_0^{t_1} \frac{1}{5} e^{-t_2/5}\, dt_2\, dt_1$$
$$= \int_0^{10} \frac{1}{10} \left[1 - e^{-t_1/5}\right] dt_1$$
$$= 1 - \frac{1}{10} \int_0^{10} e^{-t_1/5}\, dt_1$$
$$= 1 - \frac{1}{10}(5) \int_0^{10} \frac{1}{5} e^{-t_1/5}\, dt_1$$ [Rewriting using an exp PDF]

$$= 1 - \frac{1}{2}\left(1 - e^{-10/5}\right) = 1 - \frac{1}{2}\left(1 - e^{-2}\right)$$
$$= 1 - \frac{1}{2} + \frac{1}{2} e^{-2}$$
$$= \frac{1}{2} + \frac{e^{-2}}{2}$$
$$\approx 0.5 + 0.07$$
$$= 0.57$$