407 Final v2 Solutions[1]

1. Question 1

Let X and Y be independent random variables. Assume that X is uniformly distributed in $[0,2]$, and Y is uniformly distributed in $[0,4]$.

Find the density of $X+Y$.

Simplify your answer to the best of your ability.

Solution. By Proposition 6.58 in the notes, we have

$$\begin{aligned} f_{X+Y}(t) = f_X * f_Y(t) &= \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)dx \\ &= \frac{1}{2}\int_{x=0}^{x=2} \frac{1}{4} 1_{0\le t-x\le 4} dx = \frac{1}{2}\int_{x=0}^{x=2} \frac{1}{4} 1_{t-4\le x\le 4} dx \\ &= \begin{cases} \frac{1}{8}\int_{x=0}^{x=t} dx & , \text{if } 0\le t<2 \\ \frac{1}{8}\int_{x=0}^{x=2} dx & , \text{if } 2\le t<4 \\ \frac{1}{8}\int_{x=t-4}^{x=2} dx & , \text{if } 4\le t<6. \end{cases} = \begin{cases} \frac{1}{8}t & , \text{if } 0\le t<2 \\ \frac{1}{4} & , \text{if } 2\le t<4 \\ \frac{1}{8}(6-t) & , \text{if } 4\le t<6. \end{cases} \end{aligned}$$

And for any $t<0$ or $t>6$, we have $f_{X+Y}(t)=0$.

2. Question 2

Let X, Y and Z be independent random variables. Suppose X and Y are exponential random variables with parameter 1. That is,

$$f_X(x) = \begin{cases} 0 & , \text{if } x<0 \\ e^{-x} & , \text{if } x\ge 0. \end{cases} \qquad f_Y(y) = f_X(y),\ \forall\, y\in\mathbf{R}.$$

Suppose Z is uniformly distributed in $[0,1]$.

Let $W = [\max(X,Y,Z^2)]^3$. Find the density of W.

Simplify this expression to the best of your ability.

Solution. Since $X,Y,Z\ge 0$, we have $\mathbf{P}(W\le t)=0$ for all $t<0$. Let $0<t<1$. Since X,Y,Z are independent, we have

$$\begin{aligned} \mathbf{P}([\max(X,Y,Z^2)]^3\le t) &= \mathbf{P}(\max(X,Y,Z^2)\le t^{1/3}) = \mathbf{P}(X\le t^{1/3})\mathbf{P}(Y\le t^{1/3})\mathbf{P}(Z\le t^{1/6}) \\ &= (1-e^{-t^{1/3}})^2\cdot(1-t^{1/6}). \end{aligned}$$

In the case $t\ge 1$, we have $\mathbf{P}(Z\le t^{1/6})=1$, so that

$$\begin{aligned} \mathbf{P}([\max(X,Y,Z^2)]^3\le t) &= \mathbf{P}(\max(X,Y,Z^2)\le t^{1/3}) = \mathbf{P}(X\le t^{1/3})\mathbf{P}(Y\le t^{1/3}) \\ &= (1-e^{-t^{1/3}})^2. \end{aligned}$$

[1]November 22, 2020, © 2020 Steven Heilman, All Rights Reserved.

We then differentiate each case separately to get the density:

$$f_W(t) = \frac{d}{dt}\mathbf{P}(W \le t) = \begin{cases} 0 & \text{, if } t \le 0 \\ \frac{d}{dt}(1-e^{-t^{1/3}})^2 \cdot (1-t^{1/6}) & \text{, if } 0 \le t < 1 \\ \frac{d}{dt}(1-e^{-t^{1/3}})^2 & \text{, if } t \ge 1. \end{cases}$$

$$= \begin{cases} 0 & \text{, if } t \le 0 \\ 2(1-e^{-t^{1/3}})(1/3)t^{-2/3} \cdot (1-t^{1/6}) + (1-e^{-t^{1/3}})^2(-1)t^{-5/6} & \text{, if } 0 \le t < 1 \\ 2(1-e^{-t^{1/3}})(1/3)t^{-2/3} & \text{, if } t \ge 1. \end{cases}$$

3. Question 3

Give an example of a random variable X such that the following conditions all hold simultaneously for X:

- X is **NOT** a discrete random variable.
- X is **NOT** a continuous random variable.

Prove that X satisfies these properties.

Solution. Here is one such example. Suppose X has CDF F_X such that

$$F_X(t) = \begin{cases} 0 & \text{, if } t < 0 \\ t & \text{, if } 0 \le t < 1/2 \\ 1 & \text{, if } t \ge 1/2. \end{cases}$$

X is not discrete, since F_X is not piecewise constant (particularly when $0 \le t < 1/2$). Also, X is not continuous since F_X is not continuous (particularly at $t = 1/2$).

4. Question 4

Suppose you take a fair three-sided die and you roll it 90 times. For each roll of the die, it has $1/3$ probability of rolling a $1, 2$ or 3.

Let A be the event that the sum of the die rolls is greater than 240. Show that

$$\mathbf{P}(A) \le \frac{1}{120}.$$

Solution. For any $n \ge 1$, define X_n be the outcome of the nth die roll. Since the die is fair, we have $\mathbf{E}X_n = (1/3)(1+2+3) = 2$ and $\text{var}(X_n) = \mathbf{E}(X_n - 2)^2 = (1/3)(1) + 0 + (1/3)(1) = 2/3$.

Let $S := X_1 + \cdots + X_{90}$ be the sum of the die rolls. Then $\mathbf{E}S = 180$, and $\text{var}(S) = 90\text{var}(X_1) = 60$. Markov's inequality says, for any $t > 0$

$$\mathbf{P}(S > t) \le \mathbf{E}S/t = 180/t.$$

This is not helpful. Instead, we use Chebyshev's inequality. This says, for any $t > 0$,

$$\mathbf{P}(|S - 180| > t) \le t^{-2}\text{var}(S) = 60t^{-2}.$$

Choosing $t = 60$ shows that $\mathbf{P}(|S - 180| > 60) \le 1/60$. Now, using symmetry of S (interchanging the roll of 3 and 1),

$$\mathbf{P}(|S - 180| > 60) = \mathbf{P}(S < 120) + \mathbf{P}(S > 240) = 2\mathbf{P}(S > 240).$$

So,

$$2\mathbf{P}(S > 240) = \mathbf{P}(|S - 180| > 60) \le 1/60.$$

5. Question 5

Let X, Y and Z be independent random variables that are each uniformly distributed in the interval $[0, 3]$. (That is, X is uniformly distributed in $[0, 3]$, Y is uniformly distributed in $[0, 3]$, and Z is uniformly distributed in $[0, 3]$.) Compute

$$\mathbf{P}(X + Y + Z \leq 1).$$

Your final answer should be a ratio of integers.

Solution. It is given that X, Y, Z have joint density $f_{X,Y,Z}(x, y, z) = 1/27$ whenever $0 \leq x, y, z \leq 3$, and $f_{X,Y,Z} = 0$ otherwise. So, by definition of the joint density, we have

$$\begin{aligned}
\mathbf{P}(X + Y + Z \leq 1) &= \iiint_{\{x+y+z\leq 1\}} f_{X,Y,Z}(x, y, z)dxdydz = \frac{1}{27}\iiint_{\{0\leq x+y+z\leq 1\}} dxdydz \\
&= \frac{1}{27}\int_{x=0}^{1}\int_{y=0}^{y=1-x}\int_{z=0}^{z=1-x-y} dzdydx = \frac{1}{27}\int_{x=0}^{1}\int_{y=0}^{y=1-x}(1 - x - y)dydx \\
&= \frac{1}{27}\int_{x=0}^{1}[y - xy - y^2/2]_{y=0}^{y=1-x}dx = \frac{1}{27}\int_{x=0}^{1}[(1 - x) - x(1 - x) - (1 - x)^2/2]dx \\
&= \frac{1}{27}\int_{x=0}^{1}[(1 - x)^2/2]dx = \frac{1}{27}[-(1 - x)^3/6]_{x=0}^{x=1} = \frac{1}{27}(1/6) = \frac{1}{162}.
\end{aligned}$$

6. Question 6

A red cube, a green cube, and a blue cube are each put in a bowl. One cube is removed from the bowl, uniformly at random, and that cube is set on the table. Then, another cube is removed from the bowl, uniformly at random, and that cube is set on the table. In this way, two cubes have been randomly removed from the bowl.

Let R be a random variable such that $R = 1$ if the red cube is removed from the bowl and $R = 0$ otherwise.

Let G be a random variable such that $G = 1$ if the green cube is removed from the bowl and $G = 0$ otherwise.

- Compute $\mathrm{cov}(R, G)$.
- Compute $\mathrm{var}(R)$.
- Compute $\mathrm{var}(3R + 6G)$.

In all cases, simplify your answer to the best of your ability.

Solution. There are three possible outcomes, corresponding to which cube remains in the bowl, and each such outcome has probability 1/3.

- If the red cube remains in the bowl, then $R = 0$ and $G = 1$.
- If the green cube remains in the bowl, then $R = 1$ and $G = 0$.
- If the blue cube remains in the bowl, then $R = 1$ and $G = 1$.

So, $\mathbf{E}R = (0+1+1)/3 = 2/3$, $\mathbf{E}G = (1+0+1)/3 = 2/3$, and $\mathbf{E}RG = (0\cdot 1+1\cdot 0+1\cdot 1)/3 = 1/3$, and

$$\mathrm{cov}(R, G) = \mathbf{E}RG - (\mathbf{E}R)(\mathbf{E}G) = 1/3 - (2/3)^2 = 1/3 - 4/9 = -1/9.$$

Since $R = 0$ or $R = 1$, $\mathbf{E}R^2 = \mathbf{E}R = 2/3$, so that $\mathrm{var}(R) = \mathbf{E}R^2 - (\mathbf{E}R)^2 = \mathbf{E}R - (\mathbf{E}R)^2 = (2/3) - (2/3)^2 = 2/3 - 4/9 = 2/9$.

Finally, $\mathrm{var}(3R+6G) = 9\mathrm{var}(R) + 36\mathrm{var}(G) + 2\cdot 18\mathrm{cov}(R,G) = 45(2/9) + 36(-1/9) = 10-4 = 6$.

7. Question 7

Let $0 < p < 1$ with $p \neq 1/2$. Let $X_1, X_2, \ldots$ be independent identically distributed random variables such that $\mathbf{P}(X_1 = 1) = p$ and $\mathbf{P}(X_1 = -1) = 1-p$.

Let $S_0 = 0$ and for any $n \geq 1$, let $S_n = X_1 + \cdots + X_n$. Show that

$$\mathbf{P}(S_{2n} = 0 \text{ for infinitely many } n \geq 1) = 0.$$

(Hint: compute $\mathbf{P}(S_{2n} = 0)$ exactly. Then estimate $\sum_{n=0}^{\infty} \mathbf{P}(S_{2n} = 0)$. You can freely use Stirling's formula, which says that $n! \sim \sqrt{2\pi n}(n/e)^n$, i.e. $\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1$.)

Solution. Since $(S_n + n)/2$ has binomial distribution with parameters n and p, $(S_{2n}/2) + n$ has binomial distribution with parameters $2n$ and p, and $\{S_{2n} = 0\} = \{S_{2n}/2 + n = n\}$, so

$$\mathbf{P}(S_{2n} = 0) = \binom{2n}{n} p^n (1-p)^n.$$

Using Stirling's formula in the form $n! \sim \sqrt{2\pi n}(n/e)^n$, we have

$$\mathbf{P}(S_{2n} = 0) = \frac{(2n)!}{(n!)^2} p^n (1-p)^n \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{2}{n}} \frac{(2n)^{2n}}{n^{2n}} p^n (1-p)^n = \frac{1}{\sqrt{2\pi n}} (4p(1-p))^n.$$

Since $p \neq 1/2$, $4p(1-p) < 1$, so this quantity decays exponentially in n, i.e.

$$\sum_{n=0}^{\infty} \mathbf{P}(S_{2n} = 0) \leq 100 \sum_{n=1}^{\infty} (4p(1-p))^n < \infty.$$

(The sum of a geometric series is finite.) We conclude by applying the Borel-Cantelli Lemma, Proposition 7.8 in the notes.

8. Question 8

In this Exercise we show a few steps towards proving the De Moivre-Laplace Theorem. Recall that $M_X(t) = \mathbf{E}e^{tX}$ is the moment generating function of a random variable X for any $t \in \mathbf{R}$. Suppose $X_1, X_2, \ldots$ are independent identically distributed random variables with $\mathbf{P}(X_1 = 1) = \mathbf{P}(X_1 = -1) = 1/2$, so $\mathbf{E}X_1 = 0$ and $\mathrm{var}(X_1) = 1$.

- Find an explicit formula for $M_{X_1}(t)$, for any $t \in \mathbf{R}$. Simplify your answer to the best of your ability.
- Using a Taylor expansion near $t = 0$, show that $|M_{X_1}(t) - [1 + t^2/2]| \leq 100t^3$ for all t near 0. (Hint: Taylor's Theorem with error term says that

  $$f(t) = f(0) + tf'(0) + (t^2/2)f''(0) + (1/2)\int_0^t (t-s)^2 f'''(s)ds,$$

  for all t near 0. You can freely use Taylor's Theorem.)
- Find an explicit formula for $M_{(X_1+\cdots+X_n)/\sqrt{n}}(t)$, for any $t \in \mathbf{R}$. Simplify your answer to the best of your ability.

- For any $t \in \mathbf{R}$, compute
$$\lim_{n\to\infty} M_{(X_1+\cdots+X_n)/\sqrt{n}}(t).$$
Relate the answer you get to the moment generating function of a standard Gaussian random variable.

Solution. We have $M_{X_1}(t) = \mathbf{E}e^{tX_1} = (1/2)e^t + (1/2)e^{-t} = (1/2)(e^t + e^{-t})$. Using independence, we then have
$$\begin{aligned} M_{(X_1+\cdots+X_n)/\sqrt{n}}(t) &= \prod_{i=1}^{n} M_{X_i/\sqrt{n}}(t) = [M_{X_1/\sqrt{n}}(t)]^n \\ &= [M_{X_1}(t/\sqrt{n})]^n. \end{aligned}$$
The last equality used that the random variables are identically distributed. From Taylor expansion near 0, we have $|M_{X_1}(t) - [1+t^2]| \le 100t^3$ for all t near 0. So,
$$M_{(X_1+\cdots+X_n)/\sqrt{n}}(t) = (1 + t^2/[2n] + O(t^4/n^2))^n.$$
Letting $n \to \infty$, we get
$$\lim_{n\to\infty} M_{(X_1+\cdots+X_n)/\sqrt{n}}(t) = \lim_{n\to\infty} (1 + t^2/[2n])^n = e^{t^2/2}.$$
The last quantity is exactly the moment generating function of a standard Gaussian, as shown in Proposition 6.48 in the notes.