Homework problems

Problem 1

Find the MGF function of $X \sim \Gamma(n, \lambda)$.

Soltuion 1

MGF of X:

$$
\begin{aligned}
M_X(t) = \mathbb{E}\left[e^{tX}\right] = & \int\limits_{-\infty}^{\infty} e^{tx} \frac{1}{\Gamma(n)} \lambda^n x^{n-1} e^{-\lambda x} dx = \\
= & \frac{1}{\Gamma(n)} \frac{\lambda^n}{(\lambda - t)^n} \int\limits_{-\infty}^{\infty} ((\lambda - t)x)^{n-1} e^{-(\lambda - t)x} d\left((\lambda - t)x\right) = \\
= & \frac{1}{\Gamma(n)} \frac{\lambda^n}{(\lambda - t)^n} \int\limits_{-\infty}^{\infty} u^{n-1} e^{-u} d\left(u\right) = \\
= & \frac{\lambda^n}{(\lambda - t)^n}
\end{aligned}
$$

Problem 2

While running errands, you need to go to the bank, then to the post office. Let $X \sim Gamma(a, \lambda)$ be your waiting time in line at the bank, and let $Y \sim Gamma(b, \lambda)$ be your waiting time in line at the post office (with the same λ for both). Assume X and Y are independent. What is the joint distribution of $T = X + Y$ (your total wait at the bank and post office) and $W = X/(X + Y)$ (the fraction of your waiting time spent at the bank)?

Solution 2

Consider transform $(X, Y) \to (T, W)$, where $T = X + Y$ and $W = \frac{X}{X+Y}$. The inverse transform is $X = TW, Y = T(1-W)$. The Jacobian of the **direct** transform is:

$$\left|\det \frac{\partial(x,y)}{\partial(t,w)}\right| = \left|\det \begin{pmatrix} w & t \\ 1-w & -t \end{pmatrix}\right| = t$$

Then,

$$\begin{aligned} f_{T,W}(t,w) &= f_{X,Y}(tw, t(1-w)) \left|\det \frac{\partial(x,y)}{\partial(t,w)}\right| = \\ &= f_X(tw) f_Y\left(t(1-w)\right) t = \\ &= t\frac{1}{\Gamma(a)}\lambda^a (tw)^{a-1} e^{-atw} \frac{1}{\Gamma(b)}\lambda^b (t(1-w))^{b-1} e^{-bt(1-w)} = \\ &= \frac{1}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}(\lambda t)^{a+b} e^{-\lambda t}\frac{1}{t} = \\ &= \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}\right)\left(\frac{1}{\Gamma(a+b)}\lambda^{a+b} t^{a+b-1} e^{-\lambda t}\right) \end{aligned}$$

$$f_{T,W}(t,w) = \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}\right)\left(\frac{1}{\Gamma(a+b)}\lambda^{a+b} t^{a+b-1} e^{-\lambda t}\right)$$

First, it tells us that $f_{T,W}(t,w) = f_T(t) f_W(w)$, so they are independent (total time is independent of the fraction at the bank). Second, we have expression for $f_T(t)$:

$$f_T(t) = \frac{1}{\Gamma(a+b)}\lambda^{a+b} t^{a+b-1} e^{-\lambda t}$$

In which we recognize $T \sim Gamma(a+b, \lambda)$. We also have expression for $f_W(w)$:

$$f_W(w) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}$$

In which we recognize $W \sim Beta(a, b)$. This gives us expression for beta function:

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Problem 3

Use the result of previous problem to find the expectation and variance of Beta distribution.

Solution 3

Result of previous problem: for $X \sim Gamma(a, \lambda)$ and $Y \sim Gamma(b, \lambda)$, we have $T = X + Y \sim Gamma(a + b, \lambda)$ and $W = \frac{X}{X+Y} \sim Beta(a, b)$ independent.

Now,

$$\mathbb{E}\left[TW\right] = \mathbb{E}\left[T\right]\mathbb{E}\left[W\right]$$

$$\mathbb{E}\left[(X+Y)\frac{X}{X+Y}\right] = \mathbb{E}\left[X+Y\right]\mathbb{E}\left[W\right]$$

$$\mathbb{E}\left[W\right] = \frac{\mathbb{E}\left[X\right]}{\mathbb{E}\left[X+Y\right]} = \frac{a/\lambda}{a/\lambda + b/\lambda} = \frac{a}{a+b}$$

Conditional expectation

Conditional expectation: definition

We know conditional probability. Consider r.v.s X (discrete) and Z (discrete with alphabet E), then for event $X = x$

$$\mathbb{P}(X = x | Z = z) = \frac{\mathbb{P}(X = x, Z = z)}{\mathbb{P}(Z = z)}$$

A function $x \to \mathbb{P}(X = x | Z = z)$ is the conditional law of X given $Z = z$.

But what if we had not one event $Z = z$, but many events $A_i = \{Z = z_i\}$? We can still do it, by plugging a set of events $\mathcal{A} = \{A_1, \dots, A_n\}$ into function $h(\cdot)$.

What if want to account for all events with their probabilities? We can still use our function $h(\cdot)$ by plugging the random variable Z into it! $h(Z) = \mathbb{E}[X|Z]$ then will be a random variable.

Properties

- $\mathbb{E}[\alpha X_1 + \beta X_2|Z] = \alpha\mathbb{E}[X_1|Z] + \beta\mathbb{E}[X_2|Z]$
- If $X \geqslant 0$ a.s., then $\mathbb{E}[X|Z] \geqslant 0$
- $\mathbb{E}\left[\mathbb{E}[X|Z]\right] = \mathbb{E}[X]$
- For any function $h(\cdot)$ we have $\mathbb{E}[Xh(Z)|Z] = h(Z)\mathbb{E}[X|Z]$ a.s.
- If X and Z are independent, then $\mathbb{E}[X|Z] = \mathbb{E}[X]$

Example 1

A stick of length 1 is broken at point X chosen uniformly at random. Given that $X = x$, we choose another breakpoint Y uniformly at the interval $[0, x]$. Find $\mathbb{E}[Y|X]$ and its mean.

Solution 1

We have $X \sim U[0, 1]$ and $Y|X = x \sim U[0, x]$.

The expected value of $U[0, x]$ distribution is $\mathbb{E}[Y|X = x] = \frac{x}{2}$.

To get $\mathbb{E}[Y|X]$ we take this expected value and plug in our random variable, so

$$\mathbb{E}[Y|X] = \frac{X}{2}$$

To find the expectation $\mathbb{E}\left[\mathbb{E}[Y|X]\right]$, we use the property:

$$\mathbb{E}\left[\mathbb{E}[Y|X]\right] = \mathbb{E}[Y] = \mathbb{E}\left[\frac{X}{2}\right] = \frac{1}{2}\mathbb{E}[X] = \frac{1}{4}$$

Example 2

Let $Z \sim \mathcal{N}(0, 1)$ and $Y = Z^2$. Find $\mathbb{E}[Y|Z]$ and $\mathbb{E}[Z|Y]$.

Solution 2

We have $Y = h(Z) = Z^2$. Therefore

$$\mathbb{E}[Y|Z] = \mathbb{E}[h(Z)|Z] = h(Z) = Z^2$$

To get the converse $\mathbb{E}[Z|Y]$, let's first start with

$$\mathbb{E}[Z|Y=y] = \sqrt{y}\mathbb{P}(Z=\sqrt{y}|Y=y) + (-\sqrt{y})\mathbb{P}(Z=-\sqrt{y}|Y=y) = \sqrt{y}p - \sqrt{y}p$$

By plugging in r.v. Y we still get 0, so $\mathbb{E}[Z|Y] = 0$.

Example 3

Let $X_1, \ldots, X_n$ be i.i.d. r.v.s and $S_n = X_1 + \ldots + X_n$. Find $\mathbb{E}[X_1|S_n]$.

Solution 3

By symmetry,

$$\mathbb{E}[X_1|S_n] = \mathbb{E}[X_2|S_n] = \ldots = \mathbb{E}[X_n|S_n]$$

By linearity,

$$\mathbb{E}[X_1|S_n] + \mathbb{E}[X_2|S_n] + \ldots + \mathbb{E}[X_n|S_n] = \mathbb{E}[S_n|S_n] = S_n$$

So,

$$\mathbb{E}[X_1|S_n] = \frac{1}{n}S_n = \overline{X}$$

Example 4

Regression uses the following to make a prediction: $\hat{Y} = \mathbb{E}[Y|X]$. Linear regression assumes $\mathbb{E}[Y|X] = a + bX$.

1. Show that an equivalent way to express this is to write $Y = a + bX + \varepsilon$ with $\mathbb{E}[\varepsilon|X] = 0$.
2. Solve for the constants a and b

Solution 4.1

1. Let $Y = a + bX + \varepsilon$, then

$$\mathbb{E}[Y|X] = \mathbb{E}[a + bX + \varepsilon|X] = a + bX + \mathbb{E}[\varepsilon|X] = a + bX$$

2. Let $\mathbb{E}[Y|X] = a + bX$ and define $\varepsilon = Y - (a + bX)$, then $Y = a + bX + \varepsilon$ and

$$\mathbb{E}[\varepsilon|X] = \mathbb{E}[Y - (a + bX)|X] = \mathbb{E}[Y|X] - \mathbb{E}[a + bX|X] = a + bX - (a + bX)$$

Solution 4.2

1. Expectation of Y

$$\mathbb{E}[Y] = \mathbb{E}\left[\mathbb{E}[Y|X]\right] = \mathbb{E}[a + bX] = a + b\mathbb{E}[X]$$

2. ε has zero mean

$$\mathbb{E}[\varepsilon] = \mathbb{E}\left[\mathbb{E}[\varepsilon|X]\right] = \mathbb{E}[0] = 0$$

3. X and ε uncorrelated

$$\mathbb{E}[\varepsilon X] = \mathbb{E}\left[\mathbb{E}[\varepsilon X|X]\right] = \mathbb{E}X\left[\mathbb{E}[\varepsilon|X]\right] = \mathbb{E}[X \cdot 0] = 0$$

Now, take $Y = a + bX + \varepsilon$ and take covariance with X at both sides:

$$\mathrm{cov}(X, Y) = \mathrm{cov}(X, a) + \mathrm{cov}(X, bX) + \mathrm{cov}(X, \varepsilon) = b\mathbb{V}\mathrm{ar}(X)$$

Therefore,

$$b = \frac{\mathrm{cov}(X, Y)}{\mathbb{V}\mathrm{ar}(X)}$$

$$a = \mathbb{E}[Y] - \frac{\mathrm{cov}(X, Y)}{\mathbb{V}\mathrm{ar}(X)}\mathbb{E}[X]$$

Problem 5

One of two identical-looking coins is picked from a hat randomly, where one coin has probability p_1 of success and the other has probability p_2 of success. Let X be the number of successes after flipping the chosen coin n times. Find the mean of X.

Solution 5

Denote $Y \sim Be(1/2)$ the r.v. that we chose coin 1, $X_1 \sim Bi(n, p_1)$ and $X_2 \sim Bi(n, p_2)$ the number of successes for trials with different coins. Then

$$X = YX_1 + (1-Y)X_2$$

Use the tower rule:

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X|Y]\right]$$

Now we need to find

$$\mathbb{E}[X|Y=y] = ynp_1 + (1-y)np_2$$

Then,

$$\mathbb{E}[X|Y] = Ynp_1 + (1-Y)np_2$$

Finally,

$$\mathbb{E}[X] = \mathbb{E}\left[Ynp_1 + (1-Y)np_2\right] = \frac{n}{2}(p_1 + p_2)$$

Problem 6

Let $N \sim Pois(\lambda_1)$ be the number of movies released in 2022. Suppose that for every movie independently the number of tickets sold in Dolgoprudnyy is $X_i \sim Pois(\lambda_2)$. Find the mean of the total number of movie tickets sold in Dolgoprudnyy in 2022.

Solution 6

Denote $S_N = X_1 + \ldots + X_N$. We need to find $\mathbb{E}[S_N]$. Use tower rule:

$$\mathbb{E}[S_N] = \mathbb{E}\left[\mathbb{E}[S_N|N]\right]$$

$$\mathbb{E}[S_N|N=n] = \mathbb{E}[X_1 + \ldots + X_N|N=n] = \mathbb{E}[X_1 + \ldots + X_n] = n\mathbb{E}[X_1] = n\lambda_2$$

$$\mathbb{E}[S_N|N] = N\lambda_2$$

$$\mathbb{E}[S_N] = \mathbb{E}\left[N\lambda_2\right] = \lambda_1\lambda_2$$

Homework problems

Homework problems

1. Emails arrive one at a time in an inbox. Let T_n be the time at which the n-th email arrives (measured on a continuous scale from some starting point in time). Suppose that the waiting times between emails are i.i.d. $Exp(\lambda)$, i.e., $T_1, T_2 - T_1, T_3 - T_2, \ldots \sim Exp(\lambda)$. Each email is non-spam with probability p, and spam with probability $q = 1 - p$ (independently of the other emails and of the waiting times). Let X be the time at which the first non-spam email arrives (so X is a continuous r.v.

   - Find the mean of X.
   - Find the MGF of X. What famous distribution does this imply that X has?

   Hint for both parts: Let N be the number of emails until the first non-spam (including that one), and write X as a sum of N terms, then condition on N.

2. Customers arrive at a store according to a Poisson process of rate λ customers per hour. Each makes a purchase with probability p, independently. Given that a customer makes a purchase, the amount spent has mean μ (in dollars) and variance σ^2.

   - Find the mean of how much a random customer spends (note that the customer may not make a purchase).
   - Find the mean of the revenue the store obtains in an 8-hour time interval, using previous subproblem.