Student Information

Name : Hasan Ege Meyvecioğlu

ID : 2449783

Answer 1

a)

If X and Y independent, then, the following equation must be satisfied:

$$f(x,y) = f(x) * f(y)$$

However, when $x^2 + y^2 \leq 1$ it is not true since:

$$f(x,y) = \frac{1}{\pi} \neq \frac{2}{\pi} * \frac{2}{\pi} = \frac{4}{\pi^2} = f(x) * f(y), x^2 + y^2 \leq 1$$

(Marginal pdf's are calculated in part b.)

b)

We have the following formulas for marginal distributions:

$$f(x) = \int f(x,y)\, dy$$

$$f(y) = \int f(x,y)\, dx$$

We need integrals for only when $x^2 + y^2 \leq 1$ because elsewhere the $f(x,y)$ is equal to zero which makes that part of integral zero:

$$f(x) = \int_{-1}^{1} f(x,y)\, dy = \int_{-1}^{1} \frac{1}{\pi}\, dy = \frac{2}{\pi} \tag{1}$$

$$f(x) = \int_{-1}^{1} f(x,y)\, dx = \int_{-1}^{1} \frac{1}{\pi}\, dx = \frac{2}{\pi} \tag{2}$$

c)

$$E(X) = \int_{-1}^{1} x * f(x)\, dx = \int_{-1}^{1} x * \frac{2}{\pi}\, dx$$

$$= (\frac{x^2}{\pi})|_{-1}^{1} = 0$$

d)

From part c), $\mu = E(X) = 0$.

$$Var(X) = E(X-\mu)^2 = \int x^2 * f(x)\,dx - \mu^2$$
$$= \int_{-1}^{1} x^2 * (\frac{2}{\pi})\,dx - \mu^2 = \int_{-1}^{1} x^2 * (\frac{2}{\pi})\,dx$$
$$= (\frac{2x^3}{3\pi})|_{-1}^{1} = \frac{4}{3\pi}$$

Answer 2

a)

We have the following formula for pdf of the uniform distribution:

$$f(x) = \frac{1}{b-a}, a < x < b$$

Since $t_A and t_B$ are uniformly distributed over [0,100] we can use the formula:

$$f(t_A) = \frac{1}{100-0} = \frac{1}{100}, 0 < t_A < 100$$
$$f(t_B) = \frac{1}{100-0} = \frac{1}{100}, 0 < t_B < 100$$

In the question it is given that T_A and T_B are independent. Therefore we can find $f(t_A, t_B)$ easily:

$$f(t_A, t_B) = f(t_A) * f(t_B) = \frac{1}{100} * \frac{1}{100} = \frac{1}{10000}$$

And we can find the cdf by integrating the pdf over the boundaries of variables:

$$F(T_A, T_B) = \int_0^{100}\int_0^{100} f(t_A, t_B)\,dx\,dy$$
$$= \int_0^{100}\int_0^{100} \frac{1}{10000}\,dx\,dy = 1$$

b)

Assume there is a coordinate system whose axis and ordinates are t_A, t_B respectively. Consider a 100x100 square on the first region of the coordinate system. We need to find the area that is bounded by the intersection of inequalities $t_A \leq 10$(A pushes the button in first 10 second) and $t_B \geq 90$(B pushes the button in last 10 second) and the square:

$$P\{t_A \leq 10, t_B \geq 90\} = \int_{90}^{100}\int_0^{10} \frac{1}{10000}\,dt_A\,dt_B$$
$$= \frac{100}{10000} = 0.1$$

c)

Use the same coordinate system. What we are looking for is the are bounded by the square and the inequality $t_A \leq t_B + 20$ divided by pdf. The corresponding integral is:

$$
\begin{aligned}
P\{t_A \leq t_B + 20\} &= \int_0^{20}\int_0^{100} \frac{1}{10000}\,dt_B\,dt_A + \int_{20}^{100}\int_{t_A-20}^{100} \frac{1}{10000}\,dt_B\,dt_A \\
&= \frac{2000}{10000} + \frac{4800}{10000} = \frac{6800}{10000} = 0.68
\end{aligned}
$$

d)

Again we can use the 100x100 square. In order to pass the test we need this inequality to be hold:

$$
\mid t_A - t_B \mid \leq 30
$$

$$
-30 \leq t_A - t_B \leq 30
$$

Therefore our required are is the are bounded by the square region, axes and this 2 inequality:

$$
t_A - t_B \leq 30
$$

$$
t_B - t_A \leq 30
$$

Calculate this area by integration:

$$
\begin{aligned}
P\{\mid t_A - t_B \mid \leq 30\} &= \int_0^{30}\int_0^{t_A+30} \frac{1}{10000}\,dt_B\,dt_A + \int_{30}^{70}\int_{t_A-30}^{t_A+30} \frac{1}{10000}\,dt_B\,dt_A + \int_{70}^{100}\int_{t_A-30}^{100} \frac{1}{10000}\,dt_B\,dt_A \\
&= \frac{1350}{10000} + \frac{2400}{10000} + \frac{1350}{10000} = \frac{5100}{10000} = 0.51
\end{aligned}
$$

Answer 3

a)

The cdf of T can be denoted as:

$$
F_T(t) = P(T \leq t) = 1 - P(T > t)
$$

If the minimum of N variables is greater than t , than all of them are greater than t:

$$
P(T > t) = P(X_1 > t, X_2 > t, X_3 > t, ..., X_N > t)
$$

From independence rule:

$$
P(T > t) = P(X_1 > t) * P(X_2 > t) - P(X_3 >) * ... * P(X_N > t)
$$

where $P(X_i > t) = e^{-\lambda_i t}, 1 \leq i \leq N$ Using this if we multiply the probabilities we get:

$$
P(T > t) = \prod_{i=1}^{N} e^{-\lambda_i t} = e^{-t\sum_{i=1}^{N}\lambda_i}
$$

Hence the cdf is:

$$
F_T(t) = 1 - P(T > t) = 1 - e^{-t\sum_{i=1}^{N}\lambda_i}
$$

b)

We can use the cdf we have found in part a to obtain expected lifetime. Firstly, we should get the pdf by the following formula:

$$f_T(t) = \frac{\partial}{\partial t} F_T(t)$$

$$= \frac{\partial}{\partial t}(1 - e^{-t\sum_{i=1}^{10} \lambda_i}) = e^{-t\sum_{i=1}^{10} \lambda_i}$$

We can find the summation of exponential variables from the given means. For exponential distrubiton it is known that:

$$mean_i = \frac{1}{\lambda_i} = \frac{10}{n}, 1 \leq i \leq 10, i = n$$

Therefore:

$$\sum_{i=1}^{10} \lambda_i = \sum_{i=1}^{N} \frac{i}{10} = 5.5$$

So:

$$f_T(t) = e^{-5.5t}$$

We can find the expectation by integrating the following formula:

$$E(t) = \int_0^\infty t * f_T(t)\, dt$$

$$= \int_0^\infty t * e^{-5.5t}\, dt = 0.03305$$

So the expected lifetime before any of the computer fails is:

$$0.03305 \text{ years} \approx 12.1 \text{ days}$$

Answer 4

In the parts of these questions, the binomial variables and distribution can be used directly. However, since the sample we are looking for has many students ($n > 30$), we can use the normal approximation to binomial distribution and apply central limit theorem with the continuity correction.

a)

U = undergraduate students
$n = 100$
$p = 0.74$
$q = 1 - p = 0.26$
$\mu = n * p = 74$

$\sigma = \sqrt{n * p * q} = 4.386$

$$P\{70 \leq U \leq 100\} = P\{69.5 < U < 100.5\}$$
$$= P\{\frac{69.5 - 74}{4.386} < Z < \frac{100.5 - 74}{4.386}\}$$
$$= P\{-1.026 < Z < 6.042\}$$
$$= \Phi(6.042) - \Phi(-1.026) = 1 - 0.15245 = 0.848$$

b)

D = doctorate students
$n = 100$
$p = 0.1$
$q = 1 - p = 0.9$
$\mu = n * p = 10$
$\sigma = \sqrt{n * p * q} = 3$

$$P\{D \leq 5\} = P\{D < 5.5\}$$
$$= P\{Z < \frac{5.5 - 10}{3}\}$$
$$= P\{Z < -1.5\}$$
$$= \Phi(-1.5) = 0.067$$