

Q.1 (20 points) Each of 10,000 people living or working in a town may independently contract a virus with a very small probability p .

- (a) Suppose that $10000p = \ln(2) \approx 0.7$. Approximately what is the probability that *at least* one person contracted the virus?

The number X of people who contracted the virus is approximately distributed as a Poisson with $\lambda = \ln(2)$. $P(X \geq 1) = 1 - P(X = 0) \approx 1 - e^{-\lambda} \approx 0.5$.

- (b) Continue from (a). Given that at least one person has the virus, (approximately) what is the *conditional* probability that *more than* one person have the virus?

Since $P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - [e^{-\lambda} + \lambda e^{-\lambda}] \approx 0.15$, we obtain $P(X \geq 2 | X \geq 1) = \frac{P(X \geq 2)}{P(X \geq 1)} \approx 0.3$.

Q.2 (20 points) Suppose that a fair coin is tossed repeatedly until it lands heads. Let X be the number of times to toss a coin until it lands heads.

- (a) Find the frequency function $p_X(k)$ for X .

$$p_X(k) = \left(\frac{1}{2}\right)^k.$$

- (b) Compute the conditional probability that a coin is tossed k times given that there are four tails in the first four outcomes. Justify your answer.

The first four consecutive tails implies that $X \geq 5$. Recall that

$$P(X \geq k) = P((k-1) \text{ consecutive tails}) = \left(\frac{1}{2}\right)^{k-1}$$

Thus, we can calculate the conditional probability of tossing a coin k times as follows.

$$P(X = k | X \geq 5) = \frac{P(X = k)}{P(X \geq 5)} = \frac{\left(\frac{1}{2}\right)^k}{\left(\frac{1}{2}\right)^4} = \left(\frac{1}{2}\right)^{k-4}, \quad k = 5, 6, 7, \dots$$

- (c) In a renewed second attempt we toss a coin again repeatedly until it lands heads. Let Y be the number of times to toss a coin until it lands heads in the second attempt, and let $Z = \min(X, Y)$. Suppose that X and Y are independent. Find $P(Z \geq k)$ for $k = 1, 2, \dots$

We can obtain

$$P(Z \geq k) = P(X \geq k, Y \geq k) = P(X \geq k) \times P(Y \geq k) = \left(\frac{1}{2}\right)^{k-1} \times \left(\frac{1}{2}\right)^{k-1} = \left(\frac{1}{4}\right)^{k-1}$$

- (d) Find the frequency function $p_Z(k)$ for Z .

If Z is a geometric random variable with success probability p then we can find

$$P(Z \geq k) = P((k-1) \text{ consecutive failures}) = (1-p)^{k-1}$$

By the result of (c) Z must be a geometric random variable with $p = \frac{3}{4}$, and has the frequency function

$$p_Z(k) = \frac{3}{4} \left(\frac{1}{4}\right)^{k-1}.$$

Or, you can calculate

$$p_Z(k) = P(Z \geq k) - P(Z \geq k+1) = \left(\frac{1}{4}\right)^{k-1} - \left(\frac{1}{4}\right)^k = \frac{3}{4} \left(\frac{1}{4}\right)^{k-1}.$$

Q.3 (25 points) Answer the following questions. Justify your choice.

- (a) Suppose that X_1, \dots, X_m are independent and identically distributed (iid) normal random variables with parameter $(\mu = 0, \sigma^2 = 1)$. Choose the *incorrect* statement.
- (i) $\sum_{k=1}^m X_k$ has a normal distribution;
 - (ii) $\sum_{k=1}^m X_k^2$ has a normal distribution;
 - (iii) $\sum_{k=1}^m X_k^2$ has a gamma distribution.
- (ii) is incorrect.
 (i) is correct; (iii) $\sum_{k=1}^m X_k^2$ has a chi-square distribution with m degrees of freedom, which is a gamma distribution.
- (b) Suppose that X_1, \dots, X_n are iid exponential random variables with parameter $\lambda = 1/2$. Choose the *incorrect* statement.
- (i) $\sum_{k=1}^n X_k$ has a normal distribution;
 - (ii) $\sum_{k=1}^n X_k$ has a gamma distribution;
 - (iii) $\sum_{k=1}^n X_k^2$ has a chi-square distribution.
- (i) or (iii) must be considered incorrect.
 (ii) is correct since $\sum_{k=1}^n X_k$ has a gamma distribution with parameter $\alpha = n$ and $\lambda = 1/2$
 (iii) is incorrect since X_k^2 does not have a gamma distribution.
- (c) Suppose that X_1, \dots, X_n are iid random variables with density function $f(x) = e^{-x}$. Choose the *correct* statement.
- (i) $\min\{X_1, \dots, X_n\}$ has the mean value 1
 - (ii) $\min\{X_1, \dots, X_n\}$ has the mean value n
 - (iii) $\min\{X_1, \dots, X_n\}$ has the mean value $\frac{1}{n}$.
- (iii) is correct since $\min\{X_1, \dots, X_n\}$ has an exponential distribution with rate parameter $\lambda = n$; thus, the mean value is $\frac{1}{n}$.
- (d) Suppose that X_1 and X_2 are independent normal random variables with respective parameters $(\mu_1 = 2, \sigma_1 = 4)$ and $(\mu_2 = 1, \sigma_2 = 3)$. Choose the *correct* statement.
- (i) $X_1 - X_2$ has a normal distribution with $\mu = 3$ and $\sigma = 7$;
 - (ii) $X_1 - X_2$ has a normal distribution with $\mu = 3$ and $\sigma = 5$;
 - (iii) $X_1 - X_2$ has a normal distribution with $\mu = 1$ and $\sigma = 1$;
 - (iv) $X_1 - X_2$ has a normal distribution with $\mu = 1$ and $\sigma = 5$;
- (iv) is correct. It has the mean $E[X_1] - E[X_2] = 2 - 1 = 1$, and the variance $\text{Var}(X_1) + \text{Var}(X_2) = 4^2 + 3^2 = 5^2$.
- (e) Suppose that X_1 and X_2 are independent standard normal random variables. Choose the *incorrect* statement.
- (i) $X_1^2 + X_2^2$ has a normal distribution;
 - (ii) $X_1^2 + X_2^2$ has an exponential distribution;

(iii) $X_1^2 + X_2^2$ has a chi-square distribution;

(i) is incorrect.

(iii) is correct since $X_1^2 + X_2^2$ has a chi-square distribution with 2 degree of freedom, which is a gamma distribution with $\alpha = 1$ and $\lambda = 1/2$.

(ii) is correct since a gamma distribution with $\alpha = 1$ is an exponential distribution.

Q.4 (15 points) Suppose that you can choose two different types of investment, say Type A and Type B, and that you will invest \$100 in just one of them. Their returns vary according to their nature of investment. In Type A, the return at the end of term is normally distributed with parameters $\mu = 110$ and $\sigma = 10$. In Type B, the return at the end of term is normally distributed with parameters $\mu = 115$ and $\sigma = 25$. Then answer the following questions.

(a) For each type of investment, what is the average return?

They are the respective mean $\mu = 110$ and $\mu = 115$.

(b) For each type of investment, find the probability that the return is greater than \$120.

$P(X > 120) = 1 - \Phi(1) = 0.1587$ for Type A; $P(X > 120) = 1 - \Phi(0.2) = 0.4207$ for Type B.

(c) Find which type of investment has the *smaller* probability that the return is less than \$100, and justify your answer. (That is, choose the type of investment with *less* chance of losing money.)

$P(X < 100) = \Phi(-1) = 0.1587$ for Type A; $P(X < 100) = \Phi(-0.6) = 0.2743$ for Type B. Thus, Type A has less chance of losing money.

Q.5 (20 points) An electrical circuit subject to a high temperature fails at random. Let X be the time of the failure of the circuit in hours. Suppose that X is a gamma random variable, and that its mean and variance are $E[X] = 3$ and $\text{Var}(X) = 1$, respectively.

(a) Identify the parameters of distribution.

By solving $E[X] = \alpha/\lambda = 3$ and $\text{Var}(X) = \alpha/\lambda^2 = 1$, we obtain $\lambda = 3$ and $\alpha = 9$.

(b) $Y = 6X$ has a chi-square distribution. What is the degree of freedom for Y ?

$6X$ has a chi-square distribution with $n = 18$ degree of freedom since $E[Y] = 18$ and $\text{Var}(Y) = 36$.

(c) Find the probability that the circuit fails within 5.8 hours [that is, the probability $P(X \leq 5.8)$].

$P(X \leq 5.8) = P(6X \leq 34.8) = 0.99$ since we find $\chi_{0.01}^2 \approx 34.8$ when $df = 18$ at the χ^2 table.

(d) How many hours can you guarantee with 95% probability for the circuit to work continuously under the high temperature? [That is, find the value h so that $P(X \geq h) = 0.95$.]

Since $P(6X \leq 6h) = P(X \leq h) = 0.05$, we have $6h = 9.39$. Therefore, $h = 1.565$ hours.