



Subject: Probability and
Statistics - I

Chapter: Unit 2

Category: Practice questions

1. CT3 April 2006 Q5

A large portfolio of policies is such that a proportion p ($0 < p < 1$) incurred claims during the last calendar year. An investigator examines a randomly selected group of 25 policies from the portfolio.

(i) Use a Poisson approximation to the binomial distribution to calculate an approximate value for the probability that there are at most 4 policies with claims in the two cases where (a) $p = 0.1$ and (b) $p = 0.2$.

(ii) Comment briefly on the above approximations, given that the exact values of the probabilities in part (i), using the binomial distribution, are 0.9020 and 0.4207 respectively.

Ans:

(i) (a) 0.89118

(b) 0.44049

(ii) in (a) error is $0.8912 - 0.9020 = 0.0108$

in (b) error is $0.4405 - 0.4207 = 0.0198$

The approximation is valid for small p , and, as p is smaller in (a), this gives a better approximation as noted with the smaller error.

2. CT3 September 2010 Q3

Suppose that in a group of insurance policies (which are independent as regards occurrence of claims), 20% of the policies have incurred claims during the last year. An auditor is examining the policies in the group one by one in random order until two policies with claims are found.

(i) Determine the probability that exactly five policies have to be examined until two policies with claims are found.

(ii) Find the expected number of policies that have to be examined until two policies with claims are found.

Ans:

(i) 0.0819

(ii) 10

3. CT3 April 2012 Q4

Claim amounts arising under a particular type of insurance policy are modelled as having a normal distribution with standard deviation £35. They are also assumed to be independent from each other.

Calculate the probability that two randomly selected claims differ by more than £100.

Ans:

0.043

4. CT3 April 2012 Q5

Claims on a group of policies arise randomly and independently of each other through time at an average rate of 2 per month.

- (i) Calculate the probability that no claims arise in a particular month.
- (ii) Calculate the probability that more than 30 claims arise in a period of one year.

Ans:

- (i) 0.1353
- (ii) 0.0958

5. CT3 September 2012 Q5

A large portfolio consists of 20% class A policies, 50% class B policies and 30% class C policies. Ten policies are selected at random from the portfolio.

- (i) Calculate the probability that there are no policies of class A among the randomly selected ten.
- (ii) (a) Calculate the expected number of class B policies among the randomly selected ten.
(b) Calculate the probability that there are more than five class B policies among the randomly selected ten.

Ans:

- (i) 0.1074
- (ii) (a) 5
(b) 0.3770

6. CT3 September 2012 Q9

An analyst is interested in using a gamma distribution with parameters $\alpha = 2$ and $\lambda = \frac{1}{2}$, that is, with density function

$$f(x) = \frac{1}{4} x e^{-\frac{1}{2}x}, \quad 0 < x < \infty.$$

(i) (a) State the mean and standard deviation of this distribution.

(b) Hence comment briefly on its shape.

(ii) Show that the cumulative distribution function is given by

$$F(x) = 1 - \left(1 + \frac{1}{2}x\right)e^{-\frac{1}{2}x}, \quad 0 < x < \infty \quad (\text{zero otherwise}).$$

Ans:

(i) (a) mean = 4, SD = 2.8

(b) As claims are non-negative and the s.d. is quite large relative to the mean, then the distribution will be quite positively skewed.

(ii) $1 - \left(1 + \frac{1}{2}x\right)e^{-\frac{1}{2}x}$

7. CT3 April 2013 Q7

A regulator wishes to inspect a sample of an insurer's claims. The insurer estimates that 10% of policies have had one claim in the last year and no policies had more than one claim. All policies are assumed to be independent.

(i) Determine the number of policies that the regulator would expect to examine before finding 5 claims.

On inspecting the sample claims, the regulator finds that actual payments exceeded initial estimates by the following amounts:

£35 £120 £48 £200 £76

(ii) Find the mean and variance of these extra amounts.

Ans:

(i) 50

(ii) Mean = 95.8, Variance = 4454.2

8. CT3 September 2014 Q4

Consider six life policies, each on one of six independent lives. Each of four of the policies has a probability of $2/3$ of giving rise to a claim within the next five years, and each of the other two policies has a probability of $1/3$ of giving rise to a claim within the next five years. It is assumed that only one claim can arise from each policy.

- (i) Calculate the expected number of claims which will arise from the six policies within the next five years.
- (ii) Calculate the probability that exactly one claim will arise from the six policies within the next five years.
- (iii) Calculate the probability that two policies chosen at random from the six policies will both give rise to claims within the next five years.

Ans:

- (i) 3.333
- (ii) 0.0494
- (iii) 0.3037

9. CT3 April 2018 Q3

The number of minutes late that students arrive at a lecture is a random variable following an exponential distribution with mean 5 minutes.

- i) Determine the probability that a student is more than 10 minutes late to the lecture.

Twenty students arrive at the lecture independently of each other.

- ii) Determine the exact probability that fewer than two of the students are more than 10 minutes late.

Ans:

- (i) 0.1353
- (ii) 0.2255

10. CT3 September 2015 Q2

Consider the following measure of skewness for a unimodal distribution:

$$\zeta = \frac{\text{mean} - \text{mode}}{\text{standard deviation}}$$

- (i) Determine the value of ζ for a gamma distribution with parameters $\alpha = 1.6$ and $\lambda = 0.2$.
- (ii) Comment on why ζ is a suitable measure of skewness for distributions with one mode.

Ans:

- (i) 0.791
- (ii) For unimodal symmetrical distributions, the mode will coincide with the mean and therefore deviations of this measure from 0 will indicate asymmetry. Also, the measure is standardised by dividing with the standard deviation to make it scale-free.

11. CT3 April 2018 Q3

The number of minutes late that students arrive at a lecture is a random variable following an exponential distribution with mean 5 minutes.

- (i) Determine the probability that a student is more than 10 minutes late to the lecture.

Twenty students arrive at the lecture independently of each other.

- (ii) Determine the exact probability that fewer than two of the students are more than 10 minutes late.

Ans:

- (i) 0.1353
- (ii) 0.2255

12. CT3 September 2018 Q3

A sports scientist is building a statistical model to describe the number of attempts a high jump athlete will have to make until she succeeds in clearing a certain height for the first time during an indoor sports event. For this model the scientist considers a geometric distribution with probability of success p . The cumulative distribution function of the geometric distribution is given as

$$F_X(x) = 1 - (1 - p)^x, \quad x = 1, 2, 3, \dots$$

- (i) (a) State the assumptions that the scientist needs to make for considering this distribution.
 (b) Comment on the validity of the assumptions in part (i)(a).

The athlete has tried n jumps without success.

- (ii) (a) Determine the probability that the athlete will require more than x additional jumps to succeed in clearing the height.
 (b) Comment on what the answer in part (ii)(a) means for the athlete.

Ans:

- (i) (a) Needs to assume that each time the athlete tries she independently has the same probability p of passing the height, i.e. that attempts here are iid.
 (b) Given that the attempts are at the same event and on the same day, it is reasonable to assume that conditions are the same (independence) and that probability of success does not change.
- (ii) (a) $P(X > x+n \mid X > n) = (1-p)^x = P(X > x)$
 (b) The lack of success on the first n jumps is irrelevant – under this model the chances of success are not any better because there have been n attempts already.

13. CS1A April 2019 Q1

The amount of money customers spend in a single trip to the supermarket is modelled using an exponential distribution with mean €15.

- (i) Calculate the probability that a randomly selected customer spends more than €20.
 (ii) Calculate the probability that a randomly selected customer spends more than €20, given that it is known that she spends more than €15.

Ans:

- (i) 0.26360
 (ii) 0.71653

14. CS1A April 2019 Q1

A survey showed that 40% of investors invest in at least two companies in order to diversify their risk.

Calculate an approximate probability that more than 100 investors have invested in at least two companies in a random sample of 300 investors.

Ans: 0.989

15. CS1A April 2022 Q2

The number of claims arriving at an insurance company is assumed to follow a Poisson process $\{N(t)\}$ $t \geq 0$ with rate $m = 2$ per year.

- (i) State the distribution of the random variable N .
- (ii) Calculate the probability of more than two claims arriving in year 2 given that five claims arrived in year 1.
- (iii) Calculate the probability of more than two claims arriving in year 2 given that no claims arrived in year 1.
- (iv) Compare the results in parts (ii) and (iii).
- (v) Identify the distribution of the time of the n th claim, justifying your answer.

Ans:

- (i) Poisson (2)
- (ii) 0.323
- (iii) 0.323
- (iv) Both probabilities are the same as a direct consequence of the Poisson process i.e. the number of events in the time interval $(s, t]$ is independent of the number of events up to time s .
- (v) The time to the n th event in a Poisson process with rate λ is the sum of n individual inter-event times. Since the inter-event time is an $\text{Exp}(m)$ random variable, the distribution of interest is $\text{Gamma}(n, m)$.
- (vi) The CDF of an exponential distribution is given by $F(x) = 1 - e^{-mx}$. The random number can be generated by solving for x the equation $F(x) = u$. This gives $x = -\log(1-u)/m$ i.e. $x = 0.1121972$.

16. CS1A September 2022 Q2

A warranty is provided for a product worth £10,000 such that the buyer is given £8,000 if it fails in the first year, £6,000 if it fails in the second year, £4,000 if it fails in the third year, £2,000 if it fails in the fourth year and zero after that. The probability of failure in a year is 0.1. Payments are only received at the first failure. The random variable X is the number of years before the first failure occurs.

(i) Determine the distribution of the random variable X , including the value of the parameter of interest, justifying your answer and stating any assumptions.

The random variable representing the payment under the warranty is denoted by Y .

(ii) Calculate the following probabilities:

- (a) $P(Y = 8,000)$
- (b) $P(Y = 6,000)$
- (c) $P(Y = 4,000)$
- (d) $P(Y = 2,000)$
- (e) $P(Y = 0)$

(iii) Calculate the expected value of the warranty payment (Y) using your answer to part (ii).

Ans:

(i) X follows a geometric distribution with parameter p such that:

$$P(X=k)=(1-p)^{k-1}p \text{ with } p=P(\text{fail})=0.1$$

Probability p is same for all trials

- (ii) a) $P(Y = 8,000) = 0.1$
- (b) $P(Y = 6,000) = 0.09$
- (c) $P(Y = 4,000) = 0.081$
- (d) $P(Y = 2,000) = 0.0729$
- (e) $P(Y = 0) = 0.6561$

(iii) £1809.8

17. CS1A September 2023 Q1

Consider a random variable, X , with a discrete uniform distribution on the integer numbers 20, 21, ..., 79.

- (i) Determine the expected value of X .
- (ii) Determine the standard deviation of X .

Ans:

- (i) 49.5
- (ii) 17.3181

18. CS1A September 2023 Q5

X and Y are independent random variables. Let W and Z be the random variables defined as follows:

$$Z = \min(X, Y), W = \max(X, Y)$$

that is, Z gives the smaller and W the larger of the observations of X and Y.

Let F_W , F_X , F_Y and F_Z denote the cumulative distribution functions of W, X, Y and Z, respectively.

(i) Identify which one of the following expressions is correct for $F_W(s)$:

- A $F_W(s) = F_X(s) F_Y(s)$
- B $F_W(s) = F_X(s) - F_Y(s)$
- C $F_W(s) = F_Y(s) - F_X(s)$
- D $F_W(s) = \frac{F_X(s)}{F_Y(s)}$

INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES

(ii) Show that $F_Z(s) = F_X(s) + F_Y(s) - F_X(s)F_Y(s)$.

The random variable X has an exponential distribution with parameter 6, and Y, independently, also has an exponential distribution with parameter 6.

(iii) (a) Identify which one of the following expressions gives the cumulative distribution function of Z:

- A $1 - e^{12s}$
- B $1 - e^{-12s}$
- C $1 - e^{-6s}$
- D $1 - e^{-3s}$

(b) State the distribution and mean of Z, using your answer to part (iii)(a).

Ans:

- (i) A
- (ii) –
- (iii) (a) B
(b) The cumulative distribution function in (iii)(a) is the exponential distribution.
The mean of $Z = \min(X, Y)$ is $1/12$



INSTITUTE OF ACTUARIAL
& QUANTITATIVE STUDIES