INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Subject: P&S

Chapter: Unit 1

Category: Practice question 1

Chapter 2 – Descriptive statistics

1. CT3 April 2010 Q1

The mean height of the women in a large population is 1.671m while the mean height of the men in the population is 1.758m. The mean height of all the members of the population is 1.712m. Calculate the percentage of the population who are women. [2]

[Ans – 52.9%]

2. CT3 April 2010 Q8 (part i only)

For sample of 100 insurance policies the following frequency distribution gives the number of policies, f, which resulted in x claims during the last year:

x	0	1	2	3
f	76	22	1	1

(i) Calculate the sample mean, standard deviation and coefficient of skewness for these data on the number of claims per policy. [4]

[Ans – mean – 0.27, s.d – 0.529, coefficient of skewness – 2.20]

3. CT3 April 2011 Q1

The numbers of claims which have arisen in the last twelve years on 60 motor policies (continuously in force over this period) are shown (sorted) below:

0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 4 4 4 4 4 5 5 5 5 6 6 6 7

Derive:

(i) The sample median, mode and mean of the number of claims. [3]
(ii) The sample inter-quartile range of the number of claims. [2]
(iii) The sample standard deviation of the number of claims. [3]

[Total 8]

[Ans – (i) – median =2, mode = 1, mean = 2.18, (ii) – 2, (iii) – 1.84]

4. CT3 September 2011 Q1

The first 20 claims that were paid out under a group of policies were for the following amounts (in units of £1,000):

3.2 2.1 6.3 4.0 3.8 4.4 6.5 7.8 2.8 5.2 7.0 8.1 4.4 5.8 1.7 2.8 5.0 3.2 3.7 4.4

For these data $\sum x = 92.2$.

Unit 1

PRACTICE QUESTION

(i) Calculate the mean of these 20 claim amounts. [1]

The next 80 claims paid out had a mean amount of £5,025.

(ii) Calculate the mean amount for the first 100 claims. [2]

[Total 3]

[Ans – (i) – 4.61, (ii) – 4942]

5. CT3 September 2011 Q2

The claims which arose in a sample of policies of a certain class gave the following frequency distribution for the number of claims per policy in the last year:

Number of claims (x)	0	1	2	3	4 or more
Number of policies (f)	15	20	10	5	0

Calculate the third order moment about the origin for these data. [3]

[Ans – 4.7]

6. CT3 April 2012 Q1

The following 24 observations give the length of time (in hours, ordered) for which a specific fully charged laptop computer will operate on battery before requiring recharging.

1.2 1.4 1.5 1.6 1.7 1.7 1.8 1.8 1.9 1.9 2.0 2.0 2.1 2.1 2.1 2.2 2.3 2.4 2.4 2.5 3.1 3.6 3.7 4.5

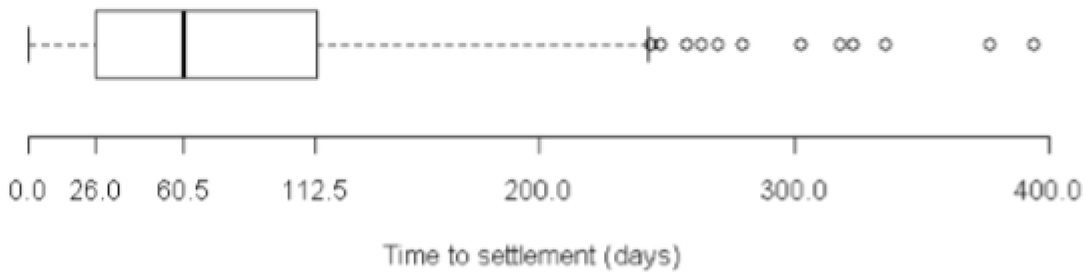
The owner of this computer is about to watch a film on his fully charged laptop.

Calculate from these data the longest showing time for a film that he can watch, so that the probability that the battery's lifetime will be sufficient for watching the entire film is 0.75. [3]

[Ans – 1.75]

7. CT3 April 2012 Q2

Data were collected on the time (in days) until each of 200 claims is settled by the insurer in a certain insurance portfolio. A boxplot of the data is shown below.



(i) Calculate the median and the interquartile range of the data using the plot. [2]

(ii) Comment on the distribution of the data as shown in the plot. [2]

[Total 4]

[Ans – (i) – 86.5]

8. CT3 September 2012 Q1

Calculate the mean, the median and the mode for the data in the following frequency table.

Observation	0	1	2	3	4
Frequency	20	54	58	28	0

[Ans – mean = 1.5875, median = 2, mode = 2]

9. CT3 September 2012 Q2

The following data are sizes of claims (ordered) for a random sample of 20 recent claims submitted to an insurance company:

174 214 264 298 335 368 381 395 402 442 487 490 564 644 686 807 1092 1328 1655 2272

(i) Calculate the interquartile range for this sample of claim sizes. [3]

(ii) Give a brief interpretation of the interquartile range calculated in part (i). [1]

[Total 4]

[Ans – (i) – 395]

10. CT3 April 2013 Q1

The following data represent the number of claims for twenty policyholders made during a year.

0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 2 2 2 3

Determine the sample mean, median, mode and standard deviation of these data. [5]

Unit 1

PRACTICE QUESTION

[Ans - mean = 0.9, median = 1, mode = 1, sd = 0.8522]

11. CT3 September 2013 Q1

The stem and leaf plot below shows 40 observations of an exchange rate.

1.21	9
1.22	4569
1.23	2679
1.24	3467889
1.25	011222345677778
1.26	00346688
1.27	
1.28	1

For these data, $\sum x = 50.000$.

(i) Find the mean, median and mode. [3]

(ii) State, with reasons, which measure of those considered in part (i) you would prefer to use to estimate the central point of the observations. [1]

[Total 4]

[Ans – (i) mean = 1.25, median = 1.252, mode = 1.257, (ii) – mean]

12. CT3 April 2014 Q1

The following sample shows the durations i x in minutes for 20 journeys from Edinburgh to Glasgow:

51 53 54 55 59 59 60 60 60 69 71 72 74 90 97 104 107 108 115 167

with $\sum_{i=1}^{20} x_i = 1585$ and $\sum_{i=1}^{20} x_i^2 = 142127$

(i) Calculate the mean and the median of this sample. [2]

(ii) Calculate the standard deviation of this sample. [2]

[Total 4]

[Ans – (i) – mean = 79.25, median = 70, (ii) – 29.48]

13. CT3 September 2014 Q1

Unit 1

PRACTICE QUESTION

A sample of marks from an exam has median 49 and interquartile range 19. The marks are rescaled by multiplying by 1.2 and adding 6.

Calculate the new median and interquartile range. [4]

[Ans – median – 64.8, IQR – 22.8]

14. CT3 April 2015 Q1

Two groups of students sat the same exam. The marks in the first group of 64 students had an average of 52 and a standard deviation of 9. The marks in the second group of 42 students had an average of 45 and a standard deviation of 8.

Calculate the average and standard deviation of the combined data set of 106 students. [4]

[Ans – average = 49.23, sd = 9.243]

15. CT3 April 2016 Q1

An university director of studies records the number of students failing the examinations of several courses. The data are presented in the following stem-and leaf plot where the stems are with units 10 and the leaves with units 1:

```
0 | 224
0 | 59
1 | 03
1 | 57889
2 |
2 |
3 | 144
3 | 5
```

(i) Determine the range of the data. [1]

(ii) Determine the median number of students failing the examinations of these courses. [1]

(iii) Determine the mean number of students failing the examinations of these courses. [1]

[Total 3]

[Ans - (i) – 33, (ii) – 16, (iii) – 16.625]

16. CT3 September 2016 Q1

Consider the following sample with 20 observations xi:

1 1 5 7 9 11 11 14 14 19 20 21 23 28 28 31 39 41 43 47

Unit 1

PRACTICE QUESTION

$$\sum_{i=1}^{20} x_i = 413 \text{ and } \sum_{i=1}^{20} x_i^2 = 12311$$

(i) Calculate the mean of this sample. [1]

(ii) Calculate the standard deviation of this sample. [2]

(iii) Calculate the median of this sample. [1]

(iv) Calculate the interquartile range of this sample. [2]

[Total 6]

[Ans – (i) – 20.65, (ii) – 14.11, (iii) – 19.5, (iv) – 19.5]

17. CT3 September 2016 Q2

A sample of data has a distribution that has a single mode and is strongly positively skewed. An analyst computes three measures of location for these data: the mean, the median and the mode.

(i) State the largest of these three location measures. [1]

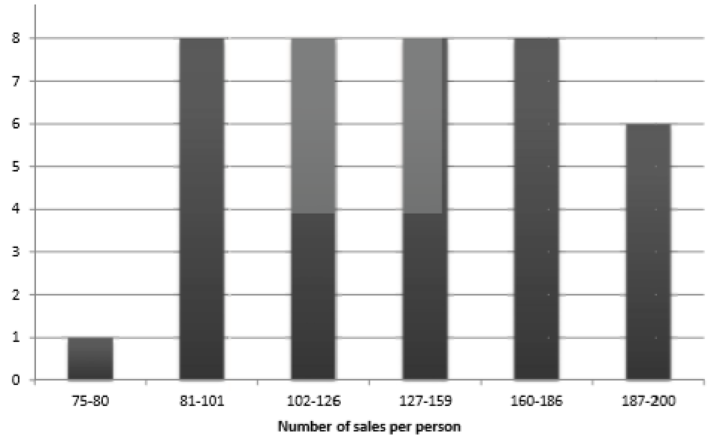
(ii) Suggest, with a reason, which would be the best measure of location. [2]

[Total 3]

[Ans – (i) – mean , (ii) – median]

18. CT3 April 2017 Q1

A company is collecting data on its sales. It has 39 sales employees and records the number of sales that each one made in a month. The data are summarised by grouping the sales employees according to the number of sales they made and presented in the following chart:



Determine the mean and standard deviation of the number of sales made by a sales employee. [4]

[Ans – mean = 138.63, sd= 37.83]

19. CT3 September 2017 Q1

The number of cans of fizzy drinks consumed by teenagers each day is the subject of an empirical study. The following data have been recorded.

Cans per day	0	1	2	3	4	5
No of teenagers	25	30	26	20	14	10

Assume that no teenager drinks more than five cans per day.

(i) Calculate the mean, median and mode for this sample. [3]

(ii) Comment on the symmetry of the observed data, using your answer to part (i) and without making any further calculations. [2]

[Total 5]

[Ans – (i) – mean = 1.984, median = 2, mode = 1]

20. CT3 April 2018 Q1

A scientist collects the following data sample on the number of plants grown on newly fertilized plots of land.

Unit 1

PRACTICE QUESTION

No of plants	1	2	3	4	5
Frequency	2	6	3	8	1

(i) Calculate the mean, median and mode of the sample. [3]

(ii) Calculate the standard deviation of the sample. [2]

[Total 5]

[Ans – (i) – mean = 3, median = 3, mode = 4,(ii) – 1.170]

21. CT3 April 2018 Q2

Consider the following data, and the corresponding sums derived from the data:

x_i :10.0   6.9   11.4   12.6   10.3   12.4   9.8

$$\sum x_i = 73.4\,,\ \sum x_i^2 = 792.22\,,\ \sum x_i^3 = 8750.972$$

(i) Determine the third moment about the mean for these data. [2]

(ii) (a) Write down the mathematical definition of the coefficient of skewness of a set of data. [1]

(b) Determine the coefficient of skewness for the data above. [2]

[Total 5]

[Ans – (i) - -4.188, (ii) b - -0.723]

22. CT3 September 2018 Q1

A data set of 20 observations has mean 45 and standard deviation 25.4. The data set is reviewed and one observation which was incorrectly recorded as 130 is now corrected to 30.

Determine the mean and standard deviation of the corrected data [4]

[Ans - mean = 40, sd = 15.825]

Chapter 3 - Fundamentals of Probability

1. CT3 April 2010 Q2

Consider a group of 10 life insurance policies, seven of which are on male lives and three of which are on female lives. Three of the 10 policies are chosen at random (one after the other, without replacement).

Find the probability that the three selected policies are all on male lives. [2]

[Ans – 0.292]

2. CT3 April 2011 Q7

An insurance company distinguishes between three types of fraudulent claims:

Type 1: legitimate claims that are slightly exaggerated

Type 2: legitimate claims that are strongly exaggerated

Type 3: false claims

Every fraudulent claim is characterised as exactly one of the three types. Assume that the probability of a newly submitted claim being a fraudulent claim of type 1 is 0.1. For type 2 this probability is 0.02, and for type 3 it is 0.003.

(i) Calculate the probability that a newly submitted claim is not fraudulent. [1]

The insurer uses a statistical software package to identify suspicious claims. If a claim is fraudulent of type 1, it is identified as suspicious by the software with probability 0.5. For a type 2 claim this probability is 0.7, and for type 3 it is 0.9.

Of all newly submitted claims, 20% are identified by the software as suspicious.

(ii) Calculate the probability that a claim that has been identified by the software as suspicious is:

(a) a fraudulent claim of type 1,

(b) a fraudulent claim of any type. [5]

(iii) Calculate the probability that a claim which has NOT been identified as suspicious by the software is in fact fraudulent. [3]

[Total 9]

[Ans – (i) – 0.877, (ii) a - 0.25, b – 0.3335, (iii) – 0.0704]

3. CT3 September 2011 Q5

Consider the random variable X taking the value X = 1 if a randomly selected person is a smoker, or X = 0 otherwise. The random variable Y describes the amount of physical exercise per week for this randomly selected person. It can take the values 0 (less than one hour of exercise per week), 1 (one to two hours) and 2 (more than two hours of exercise per week). The random variable R = (3 − Y)2(X + 1) is used as a risk index for a particular heart disease.

The joint distribution of X and Y is given by the joint probability function in the following table.

Unit 1

PRACTICE QUESTION

	y		
X	0	1	2
0	0.2	0.3	0.25
1	0.1	0.1	0.05

(i) Calculate the probability that a randomly selected person does more than two hours of exercise per week. [1]

(ii) Decide whether X and Y are independent or not and justify your answer. [2]

(iii) Derive the probability function of R. [3]

(iv) Calculate the expectation of R. [2]

[Total 8]

[Ans – (i) – 0.3, (ii) – not independent , (iv) – 5.95]

4. CT3 April 2012 Q3

Two students are selected at random without replacement from a group of 100 students, of whom 64 are male and 36 are female.

Calculate the probability that the two selected students are of different genders. [3]

[Ans – 0.465]

5. CT3 September 2013 Q7

A motor insurance company has a portfolio of 100,000 policies. It distinguishes between three groups of policyholders depending on the geographical region in which they live. The probability p of a policyholder submitting at least one claim during a year is given in the following table together with the number, n, of policyholders belonging to each group.. Each policyholder belongs to exactly one group and it is assumed that they do not move from one group to another over time.

Group	A	B	C
P	0.15	0.1	0.05
n (in 1000s)	20	20	60

It is assumed that any individual policyholder submits a claim during any year independently of claims submitted by other policyholders. It is also assumed that whether a policyholder submits any claims in a year is independent of claims in previous years conditional on belonging to a particular group.

(i) Show that the probability that a randomly selected policyholder will submit a claim in a particular year is 0.08. [2]

(ii) Calculate the probability that a randomly selected policyholder will submit a claim in a particular year given that the policyholder is not in group C. [2]

(iii) Calculate the probability for a randomly selected policyholder to belong to group A given that the policyholder submitted a claim last year. [2]

(iv) Calculate the probability that a randomly selected policyholder will submit a claim in a particular year given that the policyholder submitted a claim in the previous year. It is assumed that the insurance company does not know to which group the policyholder belongs. [3]

(v) Calculate the probability that a randomly selected policyholder will submit a claim in two consecutive years. [2]

[Total 11]

[Ans – (i) -0.08, (ii) – 0.125, (iii) – 0.375, (iv) – 0.1, (v) – 0.008]

6. CT3 April 2014 Q3

Sixty per cent of new drivers in a particular country have had additional driving education. During their first year of driving, new drivers who have not had additional driving education have a probability 0.09 of having an accident, while new drivers who have had additional driving education have a probability 0.05 of having an accident.

(a) Calculate the probability that a new driver does not have an accident during their first year of driving.

(b) Calculate the probability that a new driver has had additional driving education, given that the driver had no accidents in the first year. [5]

[Ans – (a) – 0.934, (b) – 0.610]

7. CT3 September 2014 Q2

Consider an insurer that offers two types of policy: home insurance and car insurance. 70% of all customers have a home insurance policy, and 80% of all customers have a car insurance policy. Every customer has at least one of the two types of policies.

Calculate the probability that a randomly selected customer:

(i) does not have a car insurance policy. [1]

(ii) has car insurance and home insurance. [1]

(iii) has home insurance, given that he has car insurance. [2]

(iv) does not have car insurance, given that he has home insurance. [2]

[Total 6]

Unit 1

PRACTICE QUESTION

[Ans – (i) – 20%, (ii) – 0.5, (iii) – 0.625, (iv) – 0.2857]

8. CT3 April 2016 Q4

A manufacturing company is analysing its accident record. The accidents fall into two categories:

- Minor – dealt with by first aider. Average cost £50.
- Major – hospital visit required. Average cost £1,000.

The company has 1,000 employees, of which 180 are office staff and the rest work in the factory.

The analysis shows that 10% of employees have an accident each year and 20% of accidents are major. It is assumed that an employee has no more than one accident in a year.

(i) Determine the expected total cost of accidents in a year. [2]

On further analysis it is discovered that a member of office staff has half the probability of having an accident relative to those in the factory.

(ii) Show that the probability that a given member of office staff has an accident in a year is 0.0549. [3]

(iii) Determine the probability that a randomly chosen employee who has had an accident is office staff. [2]

[Total 7]

[Ans – (i) - £24000, (ii) – 0.0549, (iii) – 0.099]

9. CT3 April 2017 Q4

An insurance company calculates car insurance premiums based on the age of the policyholder according to three age groups: Group A consists of drivers younger than 22 years old; Group B consists of drivers 22–33 years old; and, Group C consists of drivers older than 33 years.

Its portfolio consists of 10% Group A policyholders, 38% Group B policyholders and 52% Group C policyholders.

The probability of a claim in any 12-month period for a policyholder belonging to Group A, B or C is 13%, 3% and 2%, respectively.

(i) Calculate the probability that a randomly chosen policyholder from this portfolio will make a claim during a 12-month period. [3]

One of the company's policyholders has just made a claim.

(ii) Calculate the probability that the policyholder is younger than 22 years. [2]

[Total 5]

Unit 1

PRACTICE QUESTION

[Ans – (i) – 0.0348, (ii) – 0.374]

10. CT3 September 2017 Q4

An airline is analysing the punctuality of its scheduled flights. It measures departures and arrivals and classifies them as early, on time or late. From its records no flights depart early and 85% depart on time. Arrivals are early 10% of the time and late 20% of the time.

(i) Determine the probability that a flight arrives on time. [1]

Further examination shows that:

- none of the flights that depart late are early arrivals.
- 10% of the flights that depart on time are late arrivals.

(ii) Determine the probability that a flight will arrive early if its departure is on time. [3]

(iii) Show that the probability that a flight both departs on time and arrives on time is 0.665. [2]

(iv) Determine the probability that a flight will arrive on time if it departs late. [3]

(v) Determine the probability that if a flight arrives late it departed on time. [2]

[Total 11]

[Ans – (i) – 0.7, (ii) – 0.1176, (iv) – 0.233, (v) – 0.425]

11. CS1A April 2019 Q4

Alice and Bob are playing a game of dice. Two fair six-sided dice are rolled. Consider the following events:

A = ‘sum of two dice equals 3’ B = ‘sum of two dice equals 7’ C = ‘at least one of the dice shows a 1’.

(i) Show that P(C) = 11/36. [1]

(ii) Calculate P(A|C). [2]

(iii) Calculate P(B|C). [2]

(iv) Determine whether A and C are independent. [1]

(v) Determine whether B and C are independent. [1]

[Total 7]

[Ans- (ii) – 2/11, (iii) – 2/11, (iv) – not independent, (v) – not independent]

12. CS1A September 2020 Q2

A pair of fair six-sided dice is rolled once.

(i) Identify which one of the following options gives the probability that the sum of the two dice is seven:

1. 1/36
2. 1/6
3. 1/12
4. 1/3 [2]

(ii) Identify which one of the following options gives the probability that at least one dice shows three:

1. 25/36
2. 1/36
3. 11/36
4. 5/36 [2]

(iii) Identify which one of the following options gives the probability that at least one dice shows an odd number:

1. 1/4
2. 3/4
3. 1/2
4. 1/12 [2]

The random variables representing the numbers on the first and second dice are denoted by X and Y respectively.

(iv) (a) Identify which one of the following options gives the correct expression of E (X + Y | X = 4), that is the conditional expectation of the sum of the two dice given that X = 4:

1. E(Y)
2. E (X) + E(Y)
3. 4 E(X) + E(Y)
4. 4 + E(Y) [1]

(b)State a necessary assumption for deriving the answer in part (iv)(a). [1]

(c) Determine the value of E(X + Y| X = 4), using your answer to part (iv)(a). [2]

[Total 10]

[Ans – (i) – 2, (ii) – 3, (iii) – 2, (iv) a – 4, b – X and Y are independent, c – 7.5]

13. CS1A September 2020 Q3

The following data are available on three television factories that produce all the televisions used in a country.

Unit 1

PRACTICE QUESTION

Factory	% of total production	Probability of defect
A	0.35	0.020
B	0.4	0.015
C	0.25	0.010

A television is selected at random and found to have a defect (Def).

(i) Identify which one of the following expressions gives the required expression to correctly calculate the probability that the selected television was made in factory B.

$$\frac{P(\text{made in B} \mid \text{Def}) \times P(\text{Def})}{P(\text{made in A} \mid \text{Def})P(\text{Def}) + P(\text{made in B} \mid \text{Def})P(\text{Def}) + P(\text{made in C} \mid \text{Def})P(\text{Def})}$$

$$\frac{P(\text{Def} \mid \text{made in B}) \times P(\text{made in B})}{P(\text{Def} \mid \text{made in A})P(\text{made in A}) + P(\text{Def} \mid \text{made in B})P(\text{made in B}) + P(\text{Def} \mid \text{made in C})P(\text{made in C})}$$

$$\frac{P(\text{Def} \mid \text{made in B}) + P(\text{made in B})}{[P(\text{Def} \mid \text{made in A}) + P(\text{made in A})] \times [P(\text{Def} \mid \text{made in B}) + P(\text{made in B})] \times [P(\text{Def} \mid \text{made in C}) + P(\text{made in C})]}$$

$$\frac{P(\text{Def} \mid \text{made in B})}{P(\text{Def} \mid \text{made in A}) + P(\text{Def} \mid \text{made in B}) + P(\text{Def} \mid \text{made in C})}$$

[2]

(ii) Calculate, by using your answer to part (i), the probability that the selected television was produced by Manufacturer B. [2]

[Total 4]

[Ans – (i) – 2, (ii) – 0.38710]

14. CS1A September 2022 Q1

From national statistics, it is known that 7% of all drivers in a country are young drivers. It is also known that 18% of all drivers involved in road accidents are young drivers (less than 25 years old). Define the two events: A: a randomly chosen driver is involved in a road accident.

Y: a randomly chosen driver is a young driver.

(i) Determine the conditional probability P[A | Y] as a function of P[A]. [2]

(ii) Comment on the result from part (i). [1]

[Total 3]

Chapter 4 - Random Variables

Unit 1

PRACTICE QUESTION

1. CT3 April 2010 Q10

The size of claims (in units of £1,000) arising from a portfolio of house contents insurance policies can be modelled using a random variable X with probability density function (pdf) given by:

$$f_X(x) = \frac{ac^a}{x^{a+1}} \quad x \geq c$$

where a > 0 and c > 0 are the parameters of the distribution.

(i) Show that the expected value of X is E[X] =$\frac{ac}{a-1}$ for a>1

(ii) Verify that the cumulative distribution function of X is given by

$F_X(x) = 1 - (\frac{c}{x})^a$, x $\geq c$ (and = 0 for x < c). [2]

2. CT3 September 2012 Q3

Let X be a discrete random variable with the following probability distribution:

X	0	1	2	3
P(x=x)	0.4	0.3	0.2	0.1

Calculate the variance of Y, where Y = 2X + 10. [3]

[Ans – 4]

3. CT3 September 2013 Q3 (i part only)

The random variable X has a distribution with probability density function given by

$$f(x) = \begin{cases} \frac{2x}{\theta^2} & ; \quad 0 \leq x \leq \theta \\ 0 & ; \quad x < 0 \text{ or } x > \theta \end{cases}$$

where θ is the parameter of the distribution.

(i) Derive expressions in terms of θ for the expected value and the variance of X. [3]

Unit 1

PRACTICE QUESTION

[Ans – E(X) = $\frac{2\theta}{3}$, var = $\frac{\theta^2}{18}$]

4. CT3 April 2015 Q7

A continuous random variable X has the cumulative distribution function FX(x) given by:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}x^3, & 0 \le x \le 2 \\ 1, & x > 2 \end{cases}$$

(i) Determine the probability density function of X. [2]

(ii) Calculate P(0.5 <X< 1). [2]

Let . Y = $\sqrt{X}$

(iii) Determine the cumulative distribution function and the probability density function of Y. [4]

(iv) Calculate the expected values of X and Y. [4]

[Total 12]

[Ans – (i) - $\frac{3}{8}x^2$, (ii) - 0.109375, (iii) distribution f^n = $\frac{1}{8}y^6$, density f^n= $\frac{6}{8}y^5$, (iv) – E(x) =3/2 , E(Y) = 1.212183]

5. CT3 April 2016 Q3

A random variable Y has probability density function

f(y) = $\frac{\theta}{y^{\theta+1}}$ $\quad y > 1$

where θ > 0 is a parameter.

(i) Show that the probability density function of Z = ln(Y) is given by $\theta e^{-\theta z}$ and determine its range. [3]

(ii) State the distribution of Z identifying any parameters involved. [1]

[Total 4]

6. CT3 April 2017 Q2

Consider the random variable X having a distribution with probability density function:

$$f(x) = v\lambda x^{v-1} \exp exp\left(-\lambda x^{v}\right) \qquad 0 < x < \infty$$

Where ν > 0 and λ> 0 are the parameters of the distribution.

(i) Show that the cumulative distribution function of X is given by:

$F(x) = \{1 - \exp exp\left(-\lambda x^{v}\right),\ \ x > 0\ 0 \qquad x \leq 0$ [2]

You are given a value u = 0.671 from the U(0,1) distribution.

(ii) Determine by simulation a value of the random variable when ν = 1.1 and λ = 0.2. [2]

[Total 4]

[Ans – (ii) x = 4.756]

7. 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

FX(x) = 1 − (1 − p) x, x = 1, 2, 3, … (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). [3]

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. [3]

[Total 6]

[Ans – P(X> x)]

8. CS1A September 2021 Q6

Consider independent observations $y_1, y_2, \ldots, y_n$ of a random variable Y with probability density function

$$f(y) = 2cy\, e^{(-cy^2)}, \quad y > 0$$

where c > 0 is an unknown parameter. Let F(y) denote the cumulative distribution function (CDF) of Y.

(i) Identify which one of the following expressions gives the inverse of the CDF of Y

A. $y = \left\{-\frac{1}{c}\log log\,(1 - F(y))\right\}$

B. $y = 1 - \left\{-\frac{1}{c}\log log\,(1 - F(y))\right\}$

C. $y = \left\{-\frac{1}{c}\log log\,(1 - F(y))\right\}^{1/2}$

D. $y = 1 - \left\{-\frac{1}{c}\log log\,(1 - F(y))\right\}^{1/2}$ [2]

[Ans – C]

Chapter 5 - Theoretical discrete distributions

1. 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. [2]

(ii) Find the expected number of policies that have to be examined until two policies with claims are found. [1]

[Total 3]

[Ans - (i) – 0.0819, (ii) – 10]

2. CT3 September 2011 Q6

The number of claims made by each policyholder in a certain class of business is modelled as having a Poisson distribution with mean λ.

(i) Derive an expression for the probability, p, that a policyholder in this class has made at least one claim. [2]

The claims records of 20 randomly chosen policyholders were examined and the number of policyholders that made at least one claim in a year, X, was recorded.

(ii) (a) State the distribution of the random variable X and its parameters.

[Ans – (i) – 1-$e^{-\lambda}$, X – Bin(20,p)]

3. CT3 April 2012 Q5

Unit 1

PRACTICE QUESTION

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. [2]

(ii) Calculate the probability that more than 30 claims arise in a period of one year. [2]

[Total 4]

[Ans – (i) – 0.1353, (ii) – 0.0958]

4. 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. [1]

(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. [2]

[Total 3]

[Ans – (i) – 0.1074, (ii) a – 5, b – 0.3770]

5. CT3 April 2013 Q7 (i and ii part only)

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. [1]

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. [3]

[Ans – (i) – 50, (ii) mean = 95.8, variance = 4454.2]

6. CT3 September 2013 Q2

An insurance company experiences claims at a constant rate of 150 per year.

Find the approximate probability that the company receives more than 90 claims in a period of six months. [4]

Unit 1

PRACTICE QUESTION

[Ans – 0.037]

7. CT3 September 2015 Q8

Consider three groups of policyholders: A, B and C. Denote by X_A the random variable for the number of claims that a randomly chosen policyholder in group A submits during any particular calendar year. X_B and X_C denote the corresponding random variables for policyholders in groups B and C.

Assume that X_A, X_B and X_C have Poisson distributions with parameters $\lambda_A = 0.2$, $\lambda_B = 0.1$ and $\lambda_C = 0.05$, depending on the group.

Each policyholder belongs to exactly one group and group membership does not change during the lifetime of a policyholder.

Assume that:

- any individual policyholder submits a claim during any year independently of claims submitted by other policyholders.
- the number of claims a policyholder submits during a year depends on the group the policyholder belongs to, but given which group the policyholder is a member of, the number of claims submitted during a year is independent of the number of claims the policyholder submitted in previous years.

An insurance company has a portfolio of policies with 20% of policyholders belonging to group A, 20% belonging to group B and the remaining policyholders belonging to group C.

The insurance company randomly chooses one of its policyholders.

(i) Calculate the probability that this policyholder will submit at least two claims in a particular year given that he belongs to group A. [2]

Now assume that the insurance company does not know to which group the randomly selected policyholder belongs.

(ii) Show that the probability that the randomly selected policyholder submits exactly one claim in any particular year is approximately 0.0794. [3]

(iii) Determine the probability that the randomly selected policyholder belongs to group A given that the policyholder submitted exactly one claim in the previous year. [2]

(iv) Determine the probability that the randomly chosen policyholder will submit one claim during the current year given that he submitted one claim in the previous year. [5]

[Total 12]

[Ans – (i) – 0.01752, (ii) – 0.07938286 , (iii) – 0.4125, (iv) – 0.10521]

Unit 1

PRACTICE QUESTION

8. CT3 April 2017 Q8 (i and ii only)

An actuary models the number of claims X per year per policy as a discrete random variable with the following distribution

Number of claims	0	1	2	3	More than 3
Probability	*	p	p/2	p/4	p/8

where p is an unknown parameter.

(i) Show that

$P[X = 0] = \frac{8-15p}{8}$

(ii) Determine the range of possible values of p. [2]

[Ans – (ii) - p∈$[0, \frac{8}{15}]$]

9. CT3 April 2018 Q3 (part ii only)

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. [1]

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. [4]

[Total 5]

[Ans – (ii) – 0.2255]

10. CT3 September 2018 Q4

We consider three groups of policyholders: A, B and C. We denote by X_A the random variable for the number of claims that a randomly chosen policyholder in group A submits during a calendar year. X_B and X_C denote the corresponding random variables for policyholders in groups B and C. We assume that XA, XB and XC have Poisson distributions with parameters $\lambda_A = 0.2$, $\lambda_B = 0.1$ and $\lambda_C = 0.05$ depending on the group. Each policyholder belongs to exactly one group and group membership does not change during the lifetime of a policyholder. It is assumed that any individual policyholder submits claims during any year independently of claims submitted by other policyholders.

An insurance company has a portfolio of policies with 20% of policyholders belonging to group A, 20% belonging to group B and the remaining policyholders belonging to group C.

The insurance company chooses a policyholder at random.

(i) Determine the probability that this policyholder will submit at least two claims during a year given that he belongs to group A. [2]

The insurance company chooses another policyholder at random but does not know to which group he belongs.

(ii) Show that the probability this policyholder will submit exactly one claim during a year is approximately 0.0794. [3]

(iii) Calculate the probability that this policyholder belongs to group A given that he submitted exactly one claim in the previous year. [2]

[Total 7]

[Ans – (i) – 0.01752, (ii) – 0.0794, (iii) - 0.4125]

11. CS1A September 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. [3]

[Ans – 0.989]

12. 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. [2]

(ii) 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). [3]

(iii) Calculate the expected value of the warranty payment (Y) using your answer to part (ii). [2]

[Total 7]

Unit 1

PRACTICE QUESTION

[Ans – (i) - X~ geo(0.1), (ii) a - 0.1, b – 0.09, c – 0.081, d – 0.0729,e - 0.6561, (iii) – 1809.8]

Chapter 6 - Theoretical continuous distributions

1. CT3 September 2012 Q9

An analyst is interested in using a gamma distribution with parameters α = 2 and

λ = ½, that is, with density function $f(x) = \frac{1}{4}xe^{-\frac{1}{2}x} \quad 0 < x < \infty$

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

(b) Hence comment briefly on its shape. [2]

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

$F(x) = 1 - \left(1 + \frac{1}{2}x\right)e^{-\frac{1}{2}x} \quad 0 < x < 2$ (zero otherwise) [3]

[Ans – (i) a- mean = 4, sd = 2.8]

2. CT3 April 2018 Q3 (part I only)

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. [1]

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. [4]

[Total 5]

[Ans – (i) – 0.1353]

3. 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. [2]

Unit 1

PRACTICE QUESTION

(ii) Calculate the probability that a randomly selected customer spends more than €20, given that it is known that she spends more than €15. [3]

[Total 5]

[Ans – (i) - 0.26360, (ii) - 0.71653]

4. CS1A September 2020 Q9

For an empirical investigation into the amount of rent paid by tenants in a town, data on income X and rent Y have been collected. Data for a total of 300 tenants of one bedroom flats have been recorded. Assume that X and Y are both Normally distributed with expectations μ_X and μ_Y, and variances σ_X^2 and σ_Y^2. S_X and S_Y are the sample standard deviation for random samples of X and Y, respectively.

The random variable Z_X is defined as

$$Z_X = 299\frac{S_X^2}{\sigma_X^2}$$

(i) State the distribution of Z_X and all of its parameters. [2]

(ii) Write down the expectation and variance of Z_X. [2]

[Ans – (i) – Z_X ~ Chi square (299), (ii) – $E[Z_X] = 299$, $V[Z_X] = 598$]