Students' Actuarial Society

Education Committee

Probability and Statistics A - Revision Session

This course is challenging but it is possible to score well in the exam.

In this session we will try and give you more examples of some of the more basic questions that you may face in an exam. Knowing how to do these will not guarantee you will do well in the exam, but should put you in a better position to pass the course.

It is important that you understand the concepts in order to be able to apply them to scenarios and questions you will have never seen before. I know you are told this by all of your lecturers all of the time, but this course in particular is where most people get caught out. You need to understand everything rather than just learning a process or a method. The concepts in this course form the basis for courses in 3^{rd} and 4^{th} year so it is crucial you fully understand all aspects of this course and not just memorise formulas.

This session is here to hopefully reinforce the learning of some of the basics, to hopefully put everyone in the best possible position to get a pass in the course.

I have tried to target questions and areas that I have noticed come up quite often. The Probability & Statistics A exam is rarely similar 2 years in a row, so there is no guarantee that the areas covered in this revision session will appear in the exam.

Section 1

One of the main concepts of the course involves understanding distribution functions and density functions. These examples will give you practice of some of the common questions that are asked on this area. It is crucial to memorise the properties of the distribution function and density function even though they may not be explicitly asked for in the question.

Examples

1) State (without giving reasons) whether each function is a (cumulative) distribution function. For each case where the function is a distribution function, determine the corresponding density function (if it exists), the image of the distribution function and the expected value of the distribution.

(a) $F(x) = \begin{cases} 0, & if\ x < 0 \\ x^3, & if\ 0 \leq x \leq 1 \\ 1, & if\ x > 1 \end{cases}$

(b) $F(x) = \begin{cases} 1 - |x|, & if - 1 \leq x \leq 1 \\ 0, & otherwise \end{cases}$

2) Determine which of the functions are density functions. Give reasons for your answers, and, for each density, find the corresponding distribution function.

(a) $f(x) = \begin{cases} \frac{1}{2}, & if - 1 \leq x \leq 0 \\ x - 1, & if\ 1 \leq x \leq 2 \\ 0, & otherwise \end{cases}$

(b) $f(x) = \begin{cases} 1, & if - 2 \leq x \leq -1 \\ x^2, & if - 1 \leq x \leq 2 \\ 0, & otherwise \end{cases}$

(c) $f(x) = \begin{cases} 3x^2, & if\ 0 < x < 1 \\ 0, & otherwise \end{cases}$

Section 2

Another key concept is making use of density functions, both to find probability values and understanding transformations. This example will attempt to give you practice on the core types of questions which could be asked.

Example

1) Let X be a random variable with density

$$f(x) = \begin{cases} ax^2, & if\ 0 \leq x \leq 1 \\ 0, & otherwise \end{cases}$$

(a) Find a.
(b) Find the distribution function $F_x(x)$ of X
(c) Find E(X)

Now suppose Y=X^3

(d) Determine the Image of Y, ImY
(e) Find the distribution function F(y) for Y
(Hint: Express the distribution function F_Y in terms of the distribution function F_X)
(f) Find the expected value of Y

Section 3

Joint density functions are an important part of the course, and can be used to find marginal densities and then conditional densities. This is an example of finding the marginal and conditional densities before then using them to calculate conditional probabilities and expectations.

Examples

1) Let X and Y be two random variables with joint probability density function:

$$f_{X,Y}(x,y) = \begin{cases} 8xy, & 0 \le y \le x \le 1 \\ 0, & otherwise \end{cases}$$

(a) Find the marginal density function $f_x(x)$

(b) Find the marginal density function $f_Y(y)$

(c) Show that the conditional density of Y given X is given by

$$f_{Y|X}(y|x) = \frac{2y}{x^2} \qquad for\ 0 < y < x < 1$$

(d) Use $f_{Y|X}(y|x)$ to compute P(Y > ¼ | X = ¾)

(e) Use $f_{Y|X}(y|x)$ to compute E(Y | X = ¾)

2) Suppose that X ~ Exp(3) and Y ~ Exp(1) and that X and Y are independent.

(a) Determine the joint density function $f_{X,Y}$ (x, y) for X and Y

Let Z=Y/X

(b) Find the density function $f_Z(z)$

3) Suppose that X is a random variable with (marginal) density

$$f(x) = \begin{cases} 2x, & if\ 0 < x < 1 \\ 0, & otherwise \end{cases}$$

and that Y is a random variable with conditional density, given X = x,

$$f_{Y|X}(y|x) = \begin{cases} xe^{-xy} & 0 < x < 1, \quad 0 < y < \infty \\ 0, & otherwise \end{cases}$$

(a) Determine the joint density $f_{X,Y}(x,y)$ for the variables X and Y and hence, or otherwise determine P(Y > X).

(b) Without calculation, explain why $E(Y|X = x) = \frac{1}{x}$ when $0 < x < 1$

Section 4

Moment Generating Functions are a key concept of the course and will be used a lot in courses at the end of 4th year, so it is important you understand these. Questions on these will usually include showing that a certain MGF is a transformation of another MGF.

Examples

1) Suppose that $X_1, X_2, \dots \dots, X_n$ is a random variables and let

$$\bar{X} = \frac{X_1 + X_2 + \cdots \dots .. + X_n}{n}$$

denote the sample mean of $X_1, X_2, \dots .., X_n$.

(a) Show that the moment generating function, for $\bar{X}$ is given by

$$M_{\bar{X}}(\theta) = \left(M_X \left(\frac{\theta}{n} \right) \right)^n$$

where $M_X(\theta)$ is the moment generating functions for variables $X_i, i = 1,2, \dots .., n$

(b) Suppose that $X_1, X_2, \dots \dots, X_n$ are i.i.d $\Gamma(2,3)$ random variables. Determine the distribution of the sample mean $\bar{\text{X}}$. (Give reasons for your answer.)

2) Suppose that a random variable X has exponential(λ) distribution

(a) Determine the MGF $M_x(\theta)$ for X. [Show all working]

(b) Suppose that Y=cX where c is a positive constant. Show that $M_Y(\theta)$ = $M_x(c\theta)$ where M_x and M_Y are the Moment Generating Functions of X and Y respectively

(c) Hence or otherwise find $M_Y(\theta)$ and determine the distribution of Y.

Section 5

Questions like these usually appear at the end of the exam paper. These types of questions will test your ability to read through a lot of information and pick out the relevant parts required to answer the question.

Examples

1)

A speed camera is installed on a busy road and a record is kept of each time at which the camera takes a photograph of a speeding car. The times between photographs are independent and exponentially distributed with mean equal to 1/3 week.

Starting at t = 0, let X_1 denote the time until the 1st photograph is taken, let X_2 denote the time between the 1^{st} and second photograph, etc. and, in general, let X_n denote the time between the $(n - 1)^{th}$ and n^{th} photograph. (So $X_1;X_2;\ldots\ldots$ are i.i.d. exponential random variables with mean 1/3 .)

(a) Let $T_{100} = X_1 + X_2 + \ldots + X_{100}$ denote the time at which the 100^{th} photograph is taken.

i) Use a moment generating argument to determine the exact distribution of T_{100}. [4 marks]

ii) State carefully the Central Limit Theorem and use it to approximate the probability that the 100th photograph is taken sometime in the 1^{st} 35 weeks, i.e. approximate $Pr(T_{100} \leq 35$ weeks). [5 marks]

(b) It has been estimated that, given that a car has been photographed speeding, the probability that it was going more than 15 miles per hour above the speed limit is p = 0.4.

Suppose that 150 cars have been caught speeding during one year and let Y equal the number of these cars that were going more than 15 miles per hour above the speed limit. Assuming that the speed recorded for each car is independent of the speeds of all other cars, state the exact distribution of Y and use the Central Limit Theorem to determine the (approximate) probability that at least 50 cars of the 150 cars caught speeding were travelling at more than 15 miles per hour above the speed limit. [5 marks]

2)

A clerk in a small insurance company returns from holiday to discover that he has a stack of 50 claims to process. Based on past experience, the clerk estimates that the time to process each claim (in minutes) has a $\Gamma(5, 2)$ distribution independent of the process times of all other claims.

.

Let X_1 denote the time that it takes to process the first claim in the stack, let X_2 denote the time that it takes to process the second claim in the stack, etc. and, in general, let X_n denote the time that it takes to process the nth claim.

(a) State carefully the Central Limit Theorem and use it to approximate the probability that, starting at 9am, the clerk manages to process all of the claims before 11am.

(b) Suppose that, independent of all other claims, the probability, p, that a claim exceeds £1000 is 0.3. Let Y equal the number of claims, out of the 50 that the clerk must process, that are greater than £1000. State the exact distribution of Y and use the Central Limit Theorem to approximate $P(Y > 20)$