8. EXPECTATION

R.V. X
Range $\mathbb{X}$
Mass/density f_X

Seen previously EXPECTATION of X.

$$E_{f_X}[X] = \begin{cases} \sum_{x \in \mathbb{X}} x f_X(x) \\ \int_{\mathbb{X}} x f_X(x)\,dx \end{cases}$$

how the dist^n of $Y = g(X)$ can be computed.

GENERALIZED EXPECTATION

For real-valued function g defined on $\mathbb{X}$, the EXPECTATION of $g(X)$ is defined by

$$E_{f_X}[g(X)] = \begin{cases} \sum_{x \in \mathbb{X}} g(x) f_X(x) \\ \int_{\mathbb{X}} g(x) f_X(x)\,dx \end{cases}$$

(whenever the sum/integral is absolutely convergent)

i.e. the weighted average of $g(x)$ taken over $x \in \mathbb{X}$, with weights defined by f_X.

"THE LAW OF THE UNCONCIOUS STATISTICIAN"

NOTE If $Y = g(X)$ (an r.v) a fundamental result says that

$$E_{f_Y}[Y] = E_{f_X}[g(X)]$$

(see Ex 9, last Q.)

EXAMPLE

The repair time (in hours) for an engine is a continuous r.v. X where

$$X \sim Gamma(\alpha, \beta)$$

The cost of repair (in GBP) is an r.v.

$$Y = 30X + 2X^2$$

Find the expected cost of a single repair.

$$E_{f_X}[g(X)] = E_{f_X}[30X + 2X^2]$$

$$= \int_{\mathbb{X}} (30x + 2x^2) f_X(x)\,dx$$

$$= 30 \int_0^\infty x \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx$$

$$+ 2 \int_0^\infty x^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx$$

$$= \frac{30\alpha}{\beta} + \frac{2\alpha(\alpha+1)}{\beta^2}$$

$$\left[\int_0^\infty x^{\alpha+2-1} e^{-\beta x}\,dx = \frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}} \right]$$

EXAMPLE $X \sim \text{Exponential}(\lambda)$

$$g(X) = e^{3X}$$

$$E_{f_X}[g(X)] = \int_0^\infty g(x) f_X(x)\,dx$$

$$= \int_0^\infty e^{3x} \lambda e^{-\lambda x}\,dx$$

$$= \begin{cases} \dfrac{\lambda}{\lambda - 3} & (\lambda > 3) \\ \infty & \text{otherwise.} \end{cases}$$

PROPERTIES + SPECIAL CASES

(1) $g(x) = a g_1(x) + b g_2(x)$

Then

$$E_{f_X}[g(X)] = a E_{f_X}[g_1(X)] + b E_{f_X}[g_2(X)]$$

(write out integral in full, separate into two parts etc)

(2) SPECIAL CASE

$$g(x) = x^k \qquad k = 1, 2, 3, \ldots$$

$$E_{f_X}[g(X)] = E_{f_X}[X^k]$$

$$= \int x^k f_X(x)\,dx$$

– THE k^{th} MOMENT

The moments of a distribution give a more comprehensive description than the expectation

3) SPECIAL CASE

Let μ denote $E_{f_X}[X]$
Consider

$$E_{f_X}[g(X)] = E_{f_X}[(X-\mu)^k] \quad k=1,2,3,\ldots$$

- THE k^{th} CENTRAL MOMENT.

$k=2$

$$E_{f_X}[(X-\mu)^2] = \int (x-\mu)^2 f_X(x)\,dx$$

$$= \int x^2 f_X(x)\,dx - \mu^2$$

(CHECK BY WRITING OUT INTEGRAL)

Denote $E_{f_X}[(X-\mu)^2]$ by

$$Var_{f_X}[X]$$

- THE VARIANCE OF X

- measures the "spread" or "scale" of a distribution

$$Var_{f_X}[X] = \int x^2 f_X(x) - \mu^2$$

$$= E_{f_X}[X^2] - \{E_{f_X}[X]\}^2$$

$$= E_{f_X}[X(X-1)+X] - \{E_{f_X}[X]\}^2$$

(4) $g(x) = t^x \qquad t \in \mathbb{R}$

$$E_{f_X}[g(X)] = \int t^x f_X(x)\,dx$$

$$\equiv G_X(t)$$

- in the discrete case, the pgf

(5) $g(x) = e^{tx}$

$$E_{f_X}[g(X)] = \int e^{tx} f_X(x)\,dx$$

$$\equiv M_X(t) \text{ say}$$

$M_X(t)$ is the moment generating function for X

i.e. using a power series expansion of e^{tx}

$$M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!} E_{f_X}[X^r]$$

(see Exercises 8

- differentiate M_X w.r.t. t r times, evaluate at 0

$$\longrightarrow E_{f_X}[X^r]$$

$$M_X(t) = \int e^{tx} f_X(x)\,dx$$

$$G_X(t) = \int t^x f_X(x)\,dx$$

But $t^x = \exp\{x \ln t\}$

$$\therefore \quad \underline{\underline{G_X(t) = M_X(\ln t)}}$$

recall handout/notes in Ch 5 for calculations involving G_X

M_X has similar properties/uses to G_X

- each f_X has unique corresponding M_X
- if X_1, X_2 independent, $Y = X_1 + X_2$ then

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t)$$

"A KEY MGF RESULT"

<u>NOTE</u> (EXERCISES & RESULT).

$$M_X^{(r)}(t) = \frac{d^r}{ds^r}\left\{M_X(s)\right\}_{s=t}$$

$$= \int \frac{d^r}{ds^r}\left\{e^{sx}\right\}_{s=t} f_X(x)\,dx$$

$$= \int x^r e^{tx} f_X(x)\,dx$$

$$\Rightarrow M_X^{(r)}(0) = \int x^r f_X(x)\,dx$$

$$= E_{f_X}[X^r]$$

(r^{th} moment)

<u>THE CENTRAL LIMIT THEOREM</u>

Using mgfs, can prove a useful approximation theorem

– see handout

(NOT EXAMINABLE)

Implication

The "standardized" sum variable

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}}$$

is approximately Normally distributed as $n \longrightarrow \infty$.

$\therefore$ Can approximate the exact distribution of Z_n or Y_n using a Normal approximation.

EXAMPLE $Y_n \sim Binomial(n, \theta)$

$$\Rightarrow Y_n = \sum_{i=1}^{n} X_i \qquad X_i \text{ iid } Bernoulli(\theta$$

$$E_{f_{X_i}}[X_i] = \theta \qquad Var_{f_{X_i}}[X_i] = \theta(1-$$

$\Rightarrow$ By CLT, as $n \longrightarrow \infty$

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \ \dot{\sim} \ N(0,1)$$

$$\Rightarrow Y_n = \sum_{i=1}^{n} X_i \ \dot{\sim} \ N(n\theta, n\theta(1-\theta)$$

CHAPTER 9

9. JOINT DISTRIBUTIONS

Previously, modelled distributions for single r.v. X

RANGE $\mathbb{X}$
PMF / PDF f_X
CDF F_X

These ideas extend to higher dimension
$\therefore$ can model two variables jointly.

Consider two variables X, Y

$\mathbb{X}$ - range of X
$\mathbb{Y}$ - range of Y.

Let $\mathbb{X}^{(2)} = \{(x,y) : x \in \mathbb{X}, y \in \mathbb{Y}\}$

Can we describe the distribution of probability over $\mathbb{X}$?

- need a function of two variables

In the discrete case, we consider

$$f_{X,Y}(x,y) \equiv P[(X=x) \cap (Y=y)]$$
$$= P[X=x, Y=y]$$

joint mass function

- defines a table of probability values

		Y			
		1	2	3	4
X	1	.	.	.	.
	2	.	.	.	.
	3	.	.	.	.

By CONDITIONAL PROBABILITY def^n

$$P[Y=y \mid X=x] = \frac{P[X=x, Y=y]}{P[X=x]}$$

$$\therefore f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

"conditional mass function"

i.e. $f_{X,Y}(x,y) = f_X(x)\, f_{Y|X}(y|x)$

But, by the THEOREM OF TOTAL PROBABILITY

$$P[X=x] = P\Big[\bigcup_{y \in \mathcal{Y}} (X=x, Y=y)\Big]$$
$$= \sum_{y \in \mathcal{Y}} P[X=x, Y=y]$$

$$\Rightarrow f_X(x) = \sum_y f_{X,Y}(x,y)$$

MARGINAL MASS FUNCTION

$\Rightarrow$ JOINT MODEL DEFINES MARGINAL MODEL.

The random variables are independent if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

NOTE continuous case similar

JOINT PDF $f_{X,Y}(x,y)$

MARGINALS $f_X(x) = \int f_{X,Y}(x,y)\, dy$

CONDITIONAL PDF $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$

Can show that these "bivariate" functions must obey similar rules to those in the univariate case

- can extend ideas of

EXPECTATION

TRANSFORMATION

to the bivariate setting.

In particular if $X_1, \ldots, X_n$ are independent, then if

$$Y = \sum_{i=1}^{n} a_i X_i$$

$$E_{f_Y}[Y] = \sum_{i=1}^{n} a_i E_{f_{X_i}}[X_i]$$

i.e. $Y = \sum_{i=1}^{n} X_i$

$$\Rightarrow E_{f_Y}[Y] = \sum_{i=1}^{n} E_{f_{X_i}}[X_i]$$