THIRD LECTURE SUMMARY

(PROBABILITY) TREE DIAGRAM - A SPECIAL REPRESENTATION OF SAMPLE SPACE (WHEN RANDOM EXPERIMENT DONE IN STAGES) - BE ABLE TO ASSIGN ‘BRANCH’ PROBABILITIES - UNDERSTAND DIFFERENCE BETWEEN SAMPLING WITHOUT AND WITH REPLACEMENT

‘COUNTING’ FORMULAS

IN HOW MANY WAYS CAN WE:

< ARRANGE n PEOPLE IN n CHAIRS: $n!$

< SEAT r OUT OF n PEOPLE IN r CHAIRS:

$$P_{n,r} = n \times (n-1) \times (n-2) \times \ldots \quad (r \text{ FACTORS})$$

< SELECT r OUT OF n PEOPLE

$$C_{n,r} = P_{n,r} \div r!$$

RANDOM VARIABLES
AND THEIR DISTRIBUTIONS

RANDOM VARIABLE ‘TRANSLATES’ AN OUTCOME OF A RANDOM EXPERIMENT INTO A SINGLE NUMBER

A LIST OF THESE (POSSIBLE VALUES), TOGETHER WITH THE CORRESPONDING PROBABILITIES IS THE RANDOM VARIABLE’S DISTRIBUTION

SOMETIMES, IT’S SIMPLY GIVEN TO US (WITHOUT ANY EXPERIMENT), E.G.

X =	-2	-1	0	1	2
Pr:	0.13	0.19	0.23	0.27	0.18

BASED ON THIS, BE ABLE TO COMPUTE:

MEAN : = E $x \times p$ (0.18)

SIGNIFICANCE: $\bar{x}$ TENDS TO : AS THE SAMPLE SIZE INCREASES (INDEFINITELY)

STANDARD DEVIATION F = $\sqrt{(\Sigma x^2 p) - \mu^2}$

($\sqrt{1.7 - 0.18^2}$ = 1.291)

ANY PROBABILITY E.G. X > -1 (0.68)

X = 0 (0.23)

X # 1 (0.82)

-1 < X < 2 (0.50)

ETC.

‘COMMON’ DISTRIBUTIONS

BINOMIAL

EXPERIMENT: PERFORM n INDEPENDENT TRIALS, EACH WITH ONLY TWO POSSIBLE OUTCOMES (SUCCESS AND FAILURE), WITH PROBABILITIES p AND q (RESPECTIVELY)

THE (BINOMIAL) RANDOM VARIABLE X IS DEFINED AS THE TOTAL NUMBER OF SUCCESSES THUS OBTAINED

ITS (INDIVIDUAL) PROBABILITIES CAN BE COMPUTED FROM:

$$\Pr(X=i)=C_{n,i}\times p^{i}\times q^{n-i}$$

WHERE *i* RANGES FROM $0, 1, 2, \ldots$ TO n

THE CORRESPONDING MEAN AND STANDARD DEVIATION ARE COMPUTED BY $n\times p$ AND $\sqrt{n\times p\times q}$ RESPECTIVELY

WHEN USING THE PROBABILITY FORMULA, NOTE THAT $C_{n,0} = C_{n,n} = 1$ AND $p^0 = q^0 = 1$. RELATED TO THIS IS: $0\,! = 1$

POISSON

NUMBER OF CUSTOMERS, ACCIDENTS, PHONE CALLS, ETC., DURING A <u>FIXED</u> INTERVAL OF TIME (SAY t) ASSUMING THAT THE 'ARRIVAL' RATE (SAY r) IS <u>CONSTANT</u> (NO BUSY OR SLACK PERIODS)

THE CORRESPONDING DISTRIBUTION HAS ONLY ONE PARAMETER: 8 = $t \times r$

THE INDIVIDUAL PROBABILITIES ARE COMPUTED BY $\Pr(X = i) = \frac{\lambda^i}{i!} \times e^{-\lambda}$

WHERE i RANGES FROM *0* (NOT TO BE OMITTED), *1, 2, ...* TO <u>INFINITY</u>

THE CORRESPONDING MEAN AND STANDARD DEVIATION ARE: 8 AND $\sqrt{\lambda}$

IN CERTAIN CIRCUMSTANCES (p SMALL) THIS DISTRIBUTION CAN BE USED AS APPROXIMATION TO BINOMIAL (8 = $n \times p$)

SOMETIMES, INSTEAD OF TIME, THINGS MAY OCCUR IN SPACE (E.G. POTHOLES PER KM., DANDELIONS PER SQUARE METRE, ETC.). TO GET 8, WE MULTIPLY THE RATE BY LENGTH, AREA, ETC., THE REST REMAINS THE SAME

HYPERGEOMETRIC

EXPERIMENT: SAMPLING WITHOUT REPLACEMENT FROM A BOX WITH N MARBLES, K OF WHICH ARE RED

RANDOM VARIABLE: NUMBER OF RED MARBLES IN A RANDOM SAMPLE OF n

PROBABILITY OF GETTING EXACTLY i RED MARBLES IN THE SAMPLE IS:

$$\frac{C_{K,i} \times C_{N-K,n-i}}{C_{N,n}}$$

WE CAN ALSO USE IT TO COMPUTE THE PROBABILITY OF GETTING 3 SPADES (2 ACES) IN A HAND OF 5 CARDS

THE MEAN OF THE CORRESPONDING DISTRIBUTION IS $\frac{K}{N} \times n$ (SAME AS WITH REPLACEMENT), STANDARD DEVIATION IS

SLIGHTLY MORE COMPLICATED
WHEN *K* AND *N* ARE LARGE, THERE IS CLEARLY LITTLE DIFFERENCE BETWEEN THIS AND BINOMIAL DISTRIBUTION

UNIFORM

ALL PROBABILITIES ARE THE SAME (# OF DOTS WHEN ROLLING A DIE, THE 'INTEGER' DISTRIBUTION IN MINITAB)

MEAN IS AT THE CENTRE (TRUE FOR ALL SYMMETRIC DISTRIBUTIONS) - THIS YIELDS 3.5 FOR THE # OF DOTS