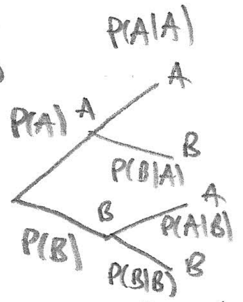
If $A_1, A_2 \ldots A_n$ are n independent events
$$P(A_1 \& A_2 \& A_3 \ldots A_n) = \prod_{i=1}^{n} P(A_i)$$

MAX $P(A_i) = 1$
MIN $P(A_i) = 0$



Tree diagram

$$P(A \& B) = P(A)P(B|A) = P(B)P(A|B)$$ conditional probability

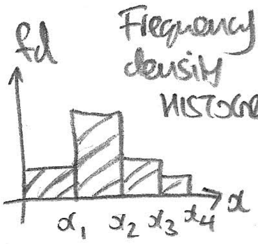
BAYES THEOREM

If $A_1 \ldots A_n$ are mutually exclusive events
$$P(A_1 \text{ OR } A_2 \text{ OR} \ldots A_n) = \sum_{i=1}^{n} P(A_i)$$

Statistics 1

Probability

See over * (Probability distributions)

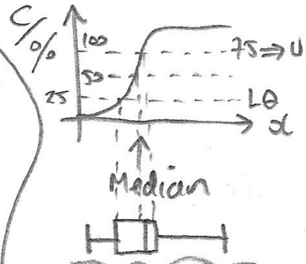


Frequency density HISTOGRAM

$$fd_j = \frac{\text{frequency}_j}{x_j - x_{j-1}} \leftarrow x_{j-1} \le x < x_j$$

Data representation

Cumulative frequency



Box & whisker

Inter-quartile range = UQ − LQ

Stem & leaf

"Back to Back"

1 3	10	6 7
2	11	1 1 3 4
6 4 4 4	12	2 2 5 9
9 8 8 8 7	13	0 0 1 2
	14	3 5

7|13 means 13.7 or 137
13|3 " 13.3 or 133
ie leaf is last digit.

Mean

$$\bar{x} = \sum_{i=1}^{n} x_i \Big/ n$$

Permutations and combinations

$$^nC_r \equiv \binom{n}{r} = \frac{n!}{(n-r)!\,r!}$$

combinations (order not important) of r objects from n objects $= {}^nC_r$

different permutations of r objects which can be made from r distinct objects is
$$^nP_r = \frac{n!}{(n-r)!}$$

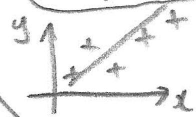
permutations of n distinct objects is $n!$

distinct permutations of n objects
- p are identical
- q of remainder are "
- r " " are "

$$= \frac{n!}{p!\,q!\,r!\ldots}$$

TRICKS
- Group objects
- Add spaces
- ⋮

Correlation and Regression*



Covariance of $\{x_i y_i\}$
$$C = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \Big/ n$$

Product moment correlation coefficient
$$r = \frac{C}{\sigma_x \sigma_y}$$
Not changed by linear transform of x, y

+1 +ve slope
−1 −ve slope
0 No correlation

* Rank correlation

Standard deviation

Standard deviation $= \sqrt{\text{VARIANCE}}$ ↘ σ^2

$$\sigma_x = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \Big/ n}$$

$\swarrow \{f_i, x_{i-1} \le x < x_i\}$

For a frequency table, estimate $\bar{x}$ and σ using mid range ie $\frac{x_i + x_{i-1}}{2}$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} \qquad \sigma_x = \sqrt{\sum_{i=1}^{n} (f_i x_i - \bar{x})^2 \Big/ \sum_{i=1}^{n} f_i}$$

↑ f_i instances of each x_i (frequency)

Spearman's rank correlation coefficient

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

Similar range $-1 \dots 0 \dots 1$ to r. (product moment).

d_i = difference between the <u>RANKS</u> for the ith item $x_i - y_i$

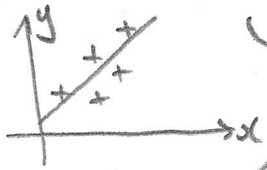
	x_i	y_i
A	5	3
B	4	2
C	2	5
D	1	4
E	3	1

Rank correlation

HARDEST QUESTIONS ARE PROBABLY COMBINATORICS — PRACTICE!

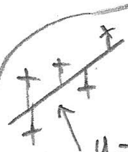
linear regression

"least squares regression" "line of best fit"



Probability of Event E $= \dfrac{\text{\# of outcomes which result in E}}{\text{total \# of possible outcomes}}$

More Statistics 1



"y on x" (vertical)

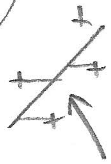
$y = a + bx$

$a = \bar{y} - b\bar{x}$

$$b = \frac{C}{\sigma_x^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})/n}{\sum_{i=1}^{n}(x_i-\bar{x})^2/n}$$

i.e. minimize $\sum_{i=1}^{n}(y_i - y)^2$

i.e. minimize $\sum_{i=1}^{n}(x - x_i)^2$



"x on y" (Horizontal)

$y = a + bx$

$a = \bar{x} - b\bar{y}$

$b = \dfrac{C}{\sigma_y^2}$

[Note $S_{xx} = \sigma_x^2 n$, $S_{yy} = \sigma_y^2 n$, $S_{xy} = Cn$]

Discrete

Probability Distributions

$X \sim D(a, b, \dots)$ "X is a random variable distributed by D, parameterised by a, b, ..."

Expectation

$E[X] = \mu = \sum_{i=1}^{n} x_i p_i$

$p_i = P(X = x_i)$

$Var[X] = \sigma^2 = \sum_{i=1}^{n}(x_i - \mu)^2 p_i = \sum_{i=1}^{n} x_i^2 p_i - \mu^2$.

BINOMIAL	$\binom{n}{x} p^x (1-p)^{n-x}$	np	$np(1-p)$	$X \sim B(n,p)$
POISSON+	$\dfrac{e^{-\lambda}\lambda^x}{x!}$	λ	λ	$X \sim P(\lambda)$
GEOMETRIC	$p(1-p)^{x-1}$	p^{-1}	$\dfrac{1-p}{p^2}$	$X \sim G(p)$
	$P(X=x)$	μ	σ	
	Probability of $X = x$	MEAN	STDEVIATION	

p is prob. of success of a trial. X = # of successes in n trials

X = # of events in given time interval

p is prob. of success of a trial. X = # of trials till first success.

+ Stats 2,3