Edexcel Further Maths AS-level

Further Statistics 2

Formula Sheet

Provided in formula book

Not provided in formula book

Linear Regression

Equation of the regression line of y on x:	$y = a + bx$	$b = \frac{S_{xy}}{S_{xx}}$, $a = \bar{y} - b\bar{x}$

Summary Statistics

For a set of n pairs of values (x_i, y_i):

$$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

$$S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

$$S_{xx} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n}$$

Product Moment Correlation Coefficient

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\{\Sigma(x_i - \bar{x})^2\}\{\Sigma(y_i - \bar{y})^2\}}} = \frac{\Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n}}{\sqrt{\left(\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right)\left(\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}\right)}}$$

Residual Sum of Squares (RSS)

$$RSS = S_{yy} - \frac{(S_{xy})^2}{S_{xx}} = S_{yy}(1 - r^2)$$

Spearman's Rank Correlation Coefficient

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

$n =$ *number of pairs of observations*
$d =$ *difference between ranks of each observation*

$r = -1$	$r = 0$	$r = +1$
Rankings are in exact reverse order	No correlation between rankings	Rankings in perfect agreement

Continuous Probability Distributions

For a continuous random variable X with probability density function $f(x)$:	$f(x) \geq 0$ for all $x \in \mathbb{R}$
	$P(a < X < b) = \int_a^b f(x)\,dx$
	$\int_{-\infty}^{+\infty} f(x)\,dx = 1$

Probability density function	$f(x) = \frac{dF(x)}{dx}$
Cumulative distribution function	$F(x_o) = P(X \leq x_o) = \int_{-\infty}^{x_o} f(x)\,dx$

Expectation (mean)	$E(X) = \mu = \int xf(x)\,dx$
Variance	$Var(X) = \sigma^2 = \int (x-\mu)^2 f(x)\,dx = \int x^2 f(x)\,dx - \mu^2$
For a function $g(X)$:	$E\big(g(X)\big) = \int g(x)f(x)\,dx$
Median, m	$\int_{-\infty}^{m} f(x)\,dx = 0.5$
Lower quartile, Q_1	$\int_{-\infty}^{Q_1} f(x)\,dx = 0.25$
Upper quartile, Q_3	$\int_{-\infty}^{Q_3} f(x)\,dx = 0.75$
n^{th} percentile, P_n	$\int_{-\infty}^{P_n} f(x)\,dx = \frac{n}{100}$
Mode	Solutions of $\frac{df(x)}{dx} = 0$ *(value at which the p.d.f is a maximum)*

$\mathrm{E}(aX+b) = a\mathrm{E}(X) + b$	$\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$

Skewness

Positive skew	$mode < median < mean$
Negative skew	$mean < median < mode$

Continuous Uniform Distribution

Probability density function:	$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & otherwise \end{cases}$

Mean	$\frac{a+b}{2}$
Variance	$\frac{(b-a)^2}{12}$
Probability distribution function	$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$