

INKMATHS FLASHCARDS

Further Statistics 2

regression, densities and tests

15 cards · fronts and backs

Print double-sided, flipping on the long edge, and each answer lands behind its question. Cut along the card borders.

Definitions in precise wording, the standard results the formulae booklet does not print, a check-yourself question per topic, and the slips that lose marks. The same cards run on the website with spaced-repetition scheduling, at inkphysics.co.uk/maths/flashcards.

STANDARD RESULT

Least squares:

$$\mathbf{b} = ?, \text{RSS} = ?$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

DEFINITION

Define: Residual.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Density and distribution:

$$\mathbf{P(a < X \leq b)} = ?, \mathbf{f(x)} = ?$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Continuous mean and variance:

$$\mathbf{E(X)} = ?, \mathbf{Var(X)} = ?$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

DEFINITION

Define: Positive skew.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Continuous uniform:

$$\mathbf{U(a, b): E(X) = ?, Var(X) = ?}$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Correlation coefficients (Spearman assumes no tied ranks):

$$\mathbf{r} = ?, \mathbf{r_s} = ?$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

RECALL

State the rules for the correlation test.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

DEFINITION

The observed y minus the y the regression line predicts at that same x , so a vertical gap and never a perpendicular one. Least squares makes the sum of their squares as small as possible, and the residuals themselves always add to zero.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$$(s_{xy})/(s_{xx}), s_{yy} - ((s_{xy})^2)/(s_{xx})$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$$\int x f(x) dx, E(X^2) - [E(X)]^2$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$$\int_a^b f(x) dx, (dF(x))/(dx)$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$$(a + b)/(2), ((b - a)^2)/(12)$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

DEFINITION

A long tail running to the right, with the bulk of the data to its left. In the single-peaked cases on this paper mode sits below median below mean, and quoting that order with the tail named earns the mark. It is evidence rather than proof: three equal averages do not force symmetry.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

RECALL

Hypotheses are about the population value ρ or ρ_s , never the sample r . Critical values fall as n grows, and the product moment version assumes a bivariate normal population.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$$(s_{xy})/(\sqrt{s_{xx} s_{yy}}), 1 - (6 \sum d^2)/(n(n^2 - 1))$$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Combining independent normals:

X, Y independent: $aX \pm bY$ is ?

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Standard error and interval, σ known:

se = ?, interval at 95 per cent = ?

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Difference of two means:

standard error = ?

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Testing a variance:

statistic = ?, degrees of freedom = ?

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

Pooled variance:

s^2 = ?, degrees of freedom = ?

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

SPOT THE ERROR

Say what is wrong with this claim:

A 95% confidence interval has a 95% chance of containing the population mean.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

CHECK YOURSELF

Samples of sizes 13 and 9 give variances 24.5 and 8.2. State the F statistic and its degrees of freedom for the alternative that the first population is the more variable.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$(\sigma)/(\sqrt{n})$, sample mean $\pm 1.96 \times se$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$N(a\mu_x \pm b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2)$, so variances always add

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$((n-1)s^2)/(\sigma^2)$ is χ^2 , $n-1$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$\sqrt{((\sigma_x^2)/(n_x) + (\sigma_y^2)/(n_y))}$. Independent samples: variances add, never standard deviations

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

SPOT THE ERROR

The population mean is fixed. Once the interval is calculated it either contains that mean or it does not, and no probability is left. What varies from sample to sample is the interval, and 95 per cent of intervals built this way would capture the mean.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

STANDARD RESULT

$((n_1-1)s_1^2 + (n_2-1)s_2^2)/(n_1 + n_2 - 2)$, $n_1 + n_2 - 2$

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS

CHECK YOURSELF

The alternative names the first population, so its variance is the numerator: $F = 24.5/8.2 = 2.99$, on 12 and 8 degrees of freedom, in that order. Larger-over-smaller is the two-tailed rule only; a one-tailed test takes its numerator from the alternative.

FURTHER STATISTICS 2: REGRESSION, DENSITIES AND TESTS