A. The big ideas

- EDA

  – Numerical and graphical summaries

  – Outliers, shape of distribution

- Parametric statistical models

  – Assume distribution family, treat only parameter as unknown

  – Data as observed values of random variables

- Repeated sampling (sampling distributions)

  – Using simulation

  – Transformations of random variables

- Checking assumptions

  – QQ (quantile) plots, etc.

  – Regression diagnostics (plot residuals against x or fitted values, etc.)

B. Kinds of (parametric) inference

- Estimation

  – Principles: method of moments

  – ... maximum likelihood

  – ... least squares

  – Performance (bias, variance, m.s.e., distributional shape)

- Confidence intervals

- Hypothesis tests

  – P-value approach

  – Critical region approach: type I and II errors, significance level, power

C. Data structures

- Single populations, simple random samples

- (Simple, linear) regression

- Comparison of populations

  - unmatched
  - matched (paired)

D. Theory

- Central limit theorem, uses, continuity correction

- Moment generating functions and their uses

- Distributions related to the Normal distribution

Tables

Tables will not be provided or allowed, or needed.

Where a question requires probability or quantile values for standard distributions, these will be given in the question, similarly to the examples on problems sheets 9 (annex) and 10.

Formulae

There will not be a formula sheet.

Questions are largely self-contained, exactly as in recent years' exams, so you don't need to memorise the table in section 2, but of course you're expected to know the definitions of the uniform, exponential and normal distributions.

R

You should be able to interpret output from **R** similar to examples seen in handouts.

You will not be asked to program in **R**.

Calculators

See Blackboard – Math Teaching – Exams – Additional Examinations information.

no graphics capability, no complex number or matrix or capability, no symbolic algebra or calculus capability, (for example, it must not be able to give you the formula $(1+x)^4 = 1 + 4x + 6x^2 + 4x^3 + x^4$, or tell you that $\sin 2A = 2\sin A \cos A$, or that the derivative of $\cos x$ is $-\sin x$, etc), no equation-solving capability, no capacity to store text, no programming capability.

Not examinable

- Quartiles (as distinct from hinges)

- MLE for non-independent data (3.11)

- Proof of central limit theorem (Rice)

- Proofs (not results) in 6.5, 6.9(b)

- Section 6.13 (explicit CI construction for Exponential distribution)

- F distribution (6.14)

- There was no section 7.7

- Confidence intervals by simulation (7.10, 7.11)