Tentative list of possible final exam problems

The final exam questions will be chosen from the final seven labs and from those on this list, which will not be 'official' until the word 'tentative' is removed from the title. The instructions on the exam include the following statement:
Your answers should be clearly expressed, using proper grammar and logical arguments. In your derivations you may use theorems from the lectures, but state or describe what they are – this is important and will affect your score.

1. [**Lecture 17**]
   (a) Define what it means for a sequence $\{X_n\}$ of random variables to converge to a constant c *in probability.*
   (b) Show that if $X_n \xrightarrow{pr} c$ and the function g is continuous at c, then $g(X_n) \xrightarrow{pr} g(c)$.

2. [**Lecture 17**] Suppose that X is a random variable taking on only positive values.
   (a) What does it mean, in terms of expectations, to say that 'X is negatively correlated with its inverse $1/X$'?
   (b) Show that in this case X is negatively correlated with $1/X$.

3. [**Lecture 18**]
   (a) Exhibit a quadratic approximation to the function $f(x) = \sqrt{x}$, valid near $x = 4$.
   (b) Use the approximation in (a) to approximate $\sqrt{5}$. Express your answer in the form $m/64$, for an integer m.

4. [**Lecture 18**] Exhibit a quadratic polynomial which furnishes a close approximation to the function $f(x) = \sin x$ when x is near $\pi/2$. Show that the error, when your polynomial is used to approximate $f(x)$, is less than or equal to $|(x - \frac{\pi}{2})^3/6|$.

5. [**Lecture 19**] A random variable X is said to have a *log-normal* distribution if $\log X$ is normally distributed: $\log X \sim N(\mu, \sigma^2)$. Suppose that $\log X$ is $N(0, 1)$. Obtain the density of X.

6. [**Lecture 19**] Suppose that the time X to completion of a certain task is exponentially distributed, and that $X_1, ..., X_n$ represent the completion times resulting from n independent repetitions of this task.

(a) If each of the completion times has mean μ, what is their density function?

(b) What is the distribution function of the random variable 'longest of these n completion times'?

(c) What is the probability that the longest of these times is less than the mean completion time?

7. [**Lecture 20**]

(a) Define what it means for a sequence $\{X_n\}$ of random variables to converge *in law* to $X \sim F$.

(b) Suppose that $X_n \sim N(\mu_n, \sigma_n^2)$ for sequences $\mu_n \to 0$ and $\sigma_n^2 \to 1$. Show that $X_n \xrightarrow{L} Z \sim N(0,1)$. [*Hint*: first show that if $\Phi(x)$ is the $N(0,1)$ distribution function, then $\Phi(\frac{x-\mu}{\sigma})$ is the $N(\mu, \sigma^2)$ distribution function.]

8. [**Lecture 20**] Suppose that a sample of values of a certain random variable X yields a sample mean $\bar{x}$ and variance s^2 with $s^2 \approx 1/\bar{x}$. Suggest, with appropriate justifications, a transformation which should stabilize the variance. [*Hint*: recall that, in the notation we used at this point, $\text{VAR}[\psi(Y)] \approx (\psi'(\mu_Y)\sigma_Y)^2$.]

9. [**Lecture 21**]

(a) Define what it means for a sequence $\{a_n\}$ of constants to converge to a finite limit a, as $n \to \infty$.

(b) Show that a bounded, increasing sequence is convergent.

10. [**Lecture 21**]

(a) State the ratio test for determining the convergence of a series $\sum a_n$ of positive terms, with partial sums $s_n = \sum_{i=1}^n a_i$.

(b) Show that all moments $E\left[X^k\right]$, $k = 0, 1, 2, ...$ of a Poisson random variable exist.

11. [**Lecture 22**]

(a) Define what it means for a sequence $\{f_n(x)\}$ of functions defined for $x \in D$ to converge *uniformly* to a function $f(x)$ on D. (We write $f_n \rightrightarrows f$ on D.)

(b) Use your answer to (a) to show that $f_n \rightrightarrows f$ on D iff the sequence $a_n = \sup_{x \in D} |f_n(x) - f(x)|$ converges to 0.

12. [**Lectures 18 and/or 22**]

(a) State Taylor's Theorem, as it applies to a function f with n derivatives throughout an interval $[a, b]$, if one wishes to approximate $f(b)$.

(b) Prove: The series $\sum_{i=0}^{\infty} \frac{x^i}{i!}$ represents the function $s(x) = e^x$.

13. [**Lecture 23**]

    (a) Define what we mean by the *probability generating function* (p.g.f.) of an integer-valued random variable (r.v.) X.

    (b) You are given that a r.v. X, taking values $n = 0, 1, 2, ...$ has the p.g.f. $E\left[z^X\right] = \frac{1}{2}\left(1 - \frac{z}{2}\right)^{-1}$. Obtain $P(X = n)$.

14. [**Lecture 23**] We have seen that one can obtain the probabilities from the probability generating function (p.g.f.), and the moments from the moment generating function (m.g.f.), by repeatedly differentiating these generating functions. Is there an analogous way to get the moments from the p.g.f.? What is it? It requires a certain assumption regarding the radius of convergence of the p.g.f. – state what this is.

15. (a) [**Lecture 23**] Define what we mean by the *cumulants* of a distribution, or of a random variable (r.v.).

    (b) [**Lecture 27**] Show that all cumulants of a normal r.v., beyond the second, are zero. [*Hint:* Recall that the moment generating function of a $N(\mu, \sigma^2)$ r.v. X is $M_X(t) = E\left[e^{tX}\right] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$.]

16. [**Lectures 23, 24**] Suppose that $X \sim bin(n, p)$.

    (a) Define what we mean by the *probability generating function* of X. Compute it.

    (b) Define what we mean by the *moment generating function* of X. Compute it.

17. [**Lectures 24, 27**] Suppose that independent r.v.s X_i are Poisson distributed with mean λ_i for $i = 1, ..., n$. What is the distribution of their sum? Justify your answer by stating relevant results from the lectures.

18. [**Lecture 25**] Prove the *Mean Value Theorem for Riemann integrals*: If f is continuous on $[a, b]$ then there is $c \in [a, b]$ for which $\int_a^b f(x)dx = f(c)(b - a)$.

19. [**Lecture 25**] Let f be a continuous function on an interval $[a, b]$ and define

$$F(x) = \int_a^x f(t)dt,$$

the indefinite integral of f. Show that F is differentiable, with derivative $F'(x) = f(x)$. Your proof will probably employ the Mean Value Theorem for Riemann integrals; if so then point out where this is used.

20. [**Lecture 26**]

   (a) Suppose that the function f is continuous and bounded on $[0, \infty)$. What does it mean to say that the integral $\int_0^\infty f(x)dx$ exists, and has the value I?

   (b) Show that the Gamma integral $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$ exists if $\alpha \geq 1$.

21. [**Lecture 26**] Define the Gamma function by $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$, for $\alpha > 0$. Assuming that the existence of this integral has already been established, show that $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

22. [**Lectures 24, 27**]

   (a) What do we mean by the *moment generating function* (m.g.f.) of a random variable X?

   (b) Describe how the moments of a random variable can be obtained from the m.g.f.

   (c) Suppose that X has the exponential density $\lambda e^{-\lambda x}$, $x > 0$. What is the m.g.f.? From this, obtain $E\left[X^k\right]$ for $k = 0, 1, 2, \ldots$.

23. [**Lectures 26, 27**]

   (a) What do we mean by the *moment generating function* of a random variable (r.v.) X?

   (b) Describe how the distribution of the sum of independent random variables can be obtained from their moment generating functions.

   (c) Recall that the moment generating function of a $N(\mu, \sigma^2)$ r.v. X is $M_X(t) = E\left[e^{tX}\right] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$. Suppose that $X_1, \ldots, X_n \overset{ind.}{\sim} N(\mu_i, \sigma_i^2)$ and $\{a_i\}$ are constants. Derive the distribution of $\sum_{i=1}^n a_i X_i$.

24. [**Lectures 26, 27**]

   (a) What do we mean by the *moment generating function* of a random variable X?

   (b) Describe how the limiting distribution of a sequence of random variables can be obtained from their moment generating functions.

   (c) Suppose that $X_n \sim bin(n, p)$ and that $p = p_n$ varies with n in such a way that $np_n \to \lambda > 0$ and $n \to \infty$. Show that $X_n \xrightarrow{L} \mathbb{P}(\lambda)$, the Poisson distribution.

25. [**Lecture 28**] Prove: If a random variable X has mean μ and variance σ^2 then $P(|X - \mu| \geq k\sigma) \leq 1/k^2$.

26. [**Lecture 28**]

   (a) State the necessary and sufficient condition under which r.v.s X and Y attain equality in the Cauchy-Schwarz Inequality. Prove that the condition is sufficient.

   (b) Using your answer to (a), show that r.v.s X and Y, with finite means and variances, are perfectly correlated if and only if Y is a linear function of X.

27. [**Lecture 28**]

   (a) State Chebyshev's Inequality, which gives a bound on the probability that a random variable X, with mean μ and variance σ^2, will be k or more standard deviations away from its mean.

   (b) State and prove the Weak Law of Large Numbers. (You may assume the result in (a).)

28. [**Lectures 17, 29**]

   (a) What does it mean to say that a statistic $\hat{\theta}_n$, computed from a sample of size n, is a *consistent estimate* of a parameter θ?

   (b) Let $X_1, ..., X_n$ be a sample from a population with mean μ and variance $\sigma^2 > 0$. Find (with appropriate justifications) a consistent estimate of the coefficient of variation σ/μ.

29. [**Lecture 29**]

   (a) What does it mean to say that "$f(x) = o(x)$ as $x \to 0$"?

   (b) Evaluate $\lim_{n\to\infty} n \log\left(1 + \frac{\sigma^2 t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)$.

30. [**Lecture 29**]

   (a) State the distribution of $\sum_{i=1}^{n} Z_i^2$, if $Z_1, ..., Z_n$ are independent $N(0,1)$ random variables.

   (b) Suppose that $X_n \sim \chi_n^2$. Show that $\frac{X_n - n}{\sqrt{n}} \xrightarrow{L} N(0, \text{VAR}[X_1])$. State the important theorem which you will no doubt use in your derivation.

31. [**Lecture 29**] Let $X_1, ..., X_n$ be independent, identically distributed random variables with moments $E\left[X^k\right] = k - 1$ for $k = 1, 2$ and 4. Define a random variable $Z_n = \left(\sum_{i=1}^{n} X_i^2 - \mu_n\right)/\sigma_n$. Show that, if the sequences $\{\mu_n\}$ and $\{\sigma_n\}$ are appropriately chosen, then $Z_n \xrightarrow{L} N(0,1)$. Exhibit the sequences $\{\mu_n\}$ and $\{\sigma_n\}$.

32. [**Lecture 29**] Let $X_1, ..., X_n$ be i.i.d., with mean μ and variance σ^2 but not necessarily normally distributed. We commonly base inferences about μ on the 't-statistic' $t = \sqrt{n}\left(\bar{X} - \mu\right)/S$, where $\bar{X}$ and S are the sample average and standard deviation. Show that this statistic has a limiting $N(0,1)$ distribution. In your derivation you may use theorems from the lectures, and results derived in the labs, but state what they are.

33. [**Lecture 30**] Suppose that the function $\mathbf{f} : \mathbf{x} \in \mathbb{R}^2 \to \mathbf{y} \in \mathbb{R}^2$ is defined by

$$\mathbf{f}(\mathbf{x}) = \mathbf{y} = \begin{pmatrix} x_1^2 \\ x_1^2 + x_2^2 \end{pmatrix},$$

and $\mathbf{g} : \mathbf{y} \in \mathbb{R}^2 \to \mathbf{z} \in \mathbb{R}^2$ is defined by

$$\mathbf{g}(\mathbf{y}) = \mathbf{z} = \begin{pmatrix} y_1/y_2 \\ y_2 \end{pmatrix}.$$

   (a) Calculate the Jacobian matrices $\left(\frac{\partial \mathbf{z}}{\partial \mathbf{y}}\right)$ and $\left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right)$.

   (b) Calculate the *determinant* of $\left(\frac{\partial \mathbf{z}}{\partial \mathbf{x}}\right)$. Express it in terms of $\mathbf{x}$. (Your solution should use (a) in such a way that no further calculations are involved.)

34. [**Lectures 30, 31**] Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be such that all mixed partial derivatives $\partial^2 f/\partial x_i \partial x_j$ exist and are continuous in a neighbourhood of a point $\mathbf{x}_0 \in D$.

   (a) Let $\mathbf{v}$ be small enough that the line from $\mathbf{x}_0$ to $\mathbf{x}_0 + \mathbf{v}$ is in D and write down a Taylor's expansion of $f(\mathbf{x}_0 + \mathbf{v})$ around $\mathbf{x}_0$ which includes the gradient and Hessian.

   (b) Suppose that $\mathbf{x}_0$ is a stationary point of f and that the Hessian is positive definite throughout a neighbourhood of $\mathbf{x}_0$. Show that then $\mathbf{x}_0$ furnishes a local minimum of f.

35. [**Lecture 31**] Let $\mathbf{A}$ be a symmetric matrix and consider the problem of maximizing a quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ over vectors satisfying $\|\mathbf{x}\| = 1$. By phrasing this as a problem involving a Lagrange multiplier, show that any such maximizing $\mathbf{x}$ is an eigenvector of $\mathbf{A}$.

36. [**Lecture 32**] The verification of the independence of the sample average and sample variance in Normal samples of size n relied on the existence of an $n \times n$ orthogonal matrix whose first row is constant. Explain how such a matrix is constructed.

37. [**Lecture 32**] The verification of the independence of the sample average and sample variance in Normal samples of size n relied on the existence of an $n \times n$ orthogonal matrix whose first row is constant. Show that if $\mathbf{H}_{n\times n}$ is such a matrix and $\mathbf{X} = (X_1, ..., X_n)'$ is a vector of sample values with sample average $\bar{X}$ and sample variance S^2, then the vector $\mathbf{Y} = \mathbf{HX}$ has first element $Y_1 = \sqrt{n}\bar{X}$ and $\sum_{i=2}^{n} Y_i^2 = (n-1)\, S^2$.

38. [**Lecture 33**] Let $X_1, ..., X_n$ be i.i.d., with common density $p(x; \boldsymbol{\theta})$.

    (a) In this context define the terms "likelihood function for $\boldsymbol{\theta}$" and "maximum likelihood estimator of $\boldsymbol{\theta}$".

    (b) We showed in class that, with probability approaching 1, the likelihood is larger at the true value of the parameter than it is at any other value. This rested on showing that, if a random variable R_n is given by

    $$R_n = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{p(X_i; \boldsymbol{\theta})}{p(X_i; \boldsymbol{\theta}_0)},$$

    then the probability of the event "$R_n > 0$" tends to 1 as $n \to \infty$. For this, we showed that R_n tends in probability to a positive constant:

    $$R_n \overset{pr}{\to} E\left[-\log \frac{p(X; \boldsymbol{\theta})}{p(X; \boldsymbol{\theta}_0)}\right] > 0.$$

    Complete the proof that $P(R_n > 0) \to 1$.

39. [**Lecture 33, 34**] Let $X_1, ..., X_n$ be i.i.d., with common density $p(x; \theta)$. (Note that θ here is a scalar, not a vector.)

    (a) Define 'Fisher information' $I(\theta)$.

    (b) We saw in class that if the right hand side of the equation $1 = \int p(x; \theta)\, dx$ can be differentiated under the integral sign, then the result is that

    $$0 = E\left[\frac{\partial}{\partial \theta} \log p(X; \theta)\right],$$

    hence that the information is $I(\theta) = E\left[\left(\frac{\partial}{\partial\theta} \log p(X; \theta)\right)^2\right]$. Assume now that the differentiation under the integral sign can be done twice, and show that

    $$I(\theta) = E\left[-\frac{\partial^2}{\partial \theta^2} \log p(X; \theta)\right].$$

40. [**Lecture 34**] Let $X_1, ..., X_n$ be a sample from a $N(\mu, \sigma^2)$ population. An unbiased estimate of μ is the sample average $\bar{X}$. Is there another with smaller variance? Why or why not?

41. [**Lecture 35**] Let $S(x)$ be a function whose minimum is being sought. (Note that here x is a scalar, not a vector.)

    (a) Write down the Newton-Raphson iteration scheme to determine the minimizing x, assuming that you have been given a starting value x_0.

    (b) Show that, if this iterative process converges, then the limit is a critical point of S.

42. [**Lecture 35**] Let $X_1, ..., X_n$ be a sample from a certain population, with a density depending on a 'location parameter' θ. This is to be estimated by an 'M-estimator', i.e. a solution to the equation $\sum_{i=1}^{n} \psi(x_i - \theta) = 0$ for a suitably chosen function ψ.

    (a) How would this equation be solved, using the Newton-Raphson method?

    (b) Suppose $\psi(t)$ is $\psi_c(t)$ for some $c > 0$, where

    $$\psi_c(t) = \begin{cases} t, & \text{if } |t| \le c, \\ sign(t) \cdot c, & \text{if } |t| \ge c. \end{cases}$$

    Describe the solutions $\hat{\theta}$ as (i) $c \to \infty$, (ii) $c \to 0$.

43. [**Lecture 36**] Recall that the Taylor series expansion of a function $\boldsymbol{\eta}(\boldsymbol{\theta}) : \mathbb{R}^p \to \mathbb{R}^n$ around a point $\boldsymbol{\theta}_0$ starts as $\boldsymbol{\eta}(\boldsymbol{\theta}) = \boldsymbol{\eta}(\boldsymbol{\theta}_0) + \mathbf{J}_\eta(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + ...$, where $\mathbf{J}_\eta(\boldsymbol{\theta})$ is the $n \times p$ Jacobian matrix. Suppose now that you are to find a minimizer of $\|\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta})\|^2$, for instance as in the nonlinear regression model $\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\theta}) + error$.

    (a) Derive the Gauss-Newton iteration scheme

    $$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \left[\mathbf{J}'_\eta(\boldsymbol{\theta}_k)\mathbf{J}_\eta(\boldsymbol{\theta}_k)\right]^{-1} \mathbf{J}'_\eta(\boldsymbol{\theta}_k)\left(\mathbf{y} - \boldsymbol{\eta}(\boldsymbol{\theta}_k)\right)$$

    for obtaining the minimizer.

    (b) Suppose that you were to compute the least squares estimator in a nonlinear regression model. Describe as explicitly as possible, in terms of regressions, residuals, etc., exactly what would be computed. Someone should be able to take your explanation and compute the solution within a standard linear regression package, without computing any matrix inverses himself.

44. [**Lecture 37**] Let $X_1, ..., X_n$ be a sample from a $N(\mu, \sigma^2)$ population.

    (a) Suppose that we know that the population mean is 0. Write down the likelihood function for σ^2.

    (b) Derive the maximum likelihood estimate of σ^2.