MATHEMATICS E-156, SPRING 2014
MATHEMATICAL FOUNDATIONS OF STATISTICAL SOFTWARE
Proof List for Final Exam

Last modified: May 11, 2014
Within each group, three of the six questions will be chosen at random.
Group 1

1. 1-1

   The variance $\mathrm{Var}[X]$ of a random variable X is defined as $E[(X - E[X])^2]$.

   Given that $E[a_1X_1 + a_2X_2] = a_1E[X_1] + a_2E[X_2]$ in all cases and that $E[X_1X_2] = E[X_1]E[X_2]$ for independent random variables, prove that

   (a) $\mathrm{Var}[X] = E[X^2] - (E[X])^2$.

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

   (c) If X_1 and X_2 are independent, $\mathrm{Var}[X_1 + X_2] = \mathrm{Var}[X_1]+\mathrm{Var}[X_2]$

2. 2-1 Variance of the mean of n independent random variables

   This is a theorem of probability.

   Suppose that $X_1, X_2, \cdots X_n$ are independent random variables, all with the same expectation μ and variance σ^2. Their mean is

   $$\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots X_n).$$

   Prove that

   $$E[\overline{X}] = \mu$$

   and that

   $$\mathrm{Var}[\overline{X}] = \frac{1}{n}\sigma^2.$$

3. 3-1

   Prove that the sum of n independent Bernoulli random variables, each with parameter p, is a binomial random variable $Y \sim \text{Binom}(n, p)$, and that

   $$E[Y] = np, \ \text{Var}\ Y = np(1-p).$$

   (a) A Bernoulli random variable X has the value 1 with probability p, 0 with probability $1-p$. Calculate its expectation and variance.

   (b) (The easy way) A binomial random variable Y is the sum of n independent Bernoulli random variables: $Y = X_1 + X_2 + \cdots + X_n$. Calculate its expectation and variance from this property alone.

   (c) If Y has the value r, then r of the X_i have the value 1, $n-r$ have the value 0. Calculate the probability of this event, which can happen in many ways, and so determine the mass (density) function $P(Y = r)$.

4. 3-2

   A contingency table with one row - why does chi-square work?

   Suppose we make n independent Bernoulli trials, each with probability p.

   A table of the results will look like this:

Success	Failure
x	$n-x$

   A table of the expected results will look like this:

Success	Failure
np	$n-np$

   The chi-square distribution with one degree of freedom has an expectation of 1. Show that the random variable that results from summing over both entries in the table also has an expectation of 1.

5. 4-1

   Poisson distribution as the limit of a binomial distribution

   Random variable X_n has a binomial distribution with parameters n and p, expectation $\lambda = np$. So $p = \lambda/n$.

   Its mass function is

   $$P(X_n = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} \frac{\lambda^x}{n^x} (1 - \frac{\lambda}{n})^{n-x}.$$

   Take the limit as $n \to \infty$ to get the mass function for a Poisson random variable X:

   $$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

6. 4-3

   Prove that if X_1 and X_2 are independent Poisson random variables with parameters λ_1 and λ_2 respectively, then $X_1 + X_2$ is Poisson with parameter $\lambda_1 + \lambda_2$.

Group 2

7. 5-2 Density function for the maximum

   Suppose that $X_1, X_2, \cdots X_n$ are independent, identically distributed continuous random variables with density function f and distribution function F.

   Let $X_{max} = \max\{X_1, X_2, \cdots X_n\}$.

   Prove that the density function for X_{max} is $f_{max}(x) = nF(x)^{n-1}f(x)$.

   Specialize to the case where the X_i are random variables from Unif$[0, \beta]$

8. 5-3

   For random variable X, define the moment generating function

   $$M(t) = E[e^{tX}].$$

   Prove that

   - The nth derivative of $M(t)$, evaluated at $t = 0$, is equal to the nth moment $E[X^n]$.
   - If X_1 and X_2 are independent random variables with moment generating functions $M_1(t)$ and $M_2(t)$, then

     $$M_{X_1+X_2}(t) = M_1(t)M_2(t)$$

9. 6-1 Confidence intervals as random variables

   Suppose that we are drawing samples of size n from a known population with mean μ and variance σ^2. The sample mean $\overline{X}$ is a random variable whose expectation is also μ. Let α be a smallish number, typically $\alpha = 0.05$. Then a "$1 - \alpha$ confidence interval" is specified by two random variables L and U with the property that

   $$P(L \geq \mu) = P(U \leq \mu) = \alpha/2.$$

   Thus the probability of the event $L < \mu < U$ is $1 - \alpha$ (typically 95%).

   Let q_1 and q_2 denote the $\alpha/2$ and $(1 - \alpha/2)$ quantiles for the sampling distribution of X.

   (a) Show that $U = \overline{X} + \mu - q_1$ and that $L = \overline{X} + \mu - q_2$. Explain why it is reasonable for U to depend on the "lower" quantile q_1 and vice versa.

   (b) Show that if the sampling distribution is symmetrical about μ, then $U = \overline{X} + q_2 - \mu$ and that $L = \overline{X} - (\mu - q_1)$.

   (c) For the normal distribution $N(0, 1)$ and $\alpha = .05$, $\mu = 0$, $q_1 = -1.96$ and $q_2 = 1.96$. Show that if the central limit theorem applies, then $L = \overline{X} - 1.96\sigma/\sqrt{n}; U = \overline{X} + 1.96\sigma/\sqrt{n}$.

10. 6-2

   A normal random variable $X \sim N(\mu, \sigma^2)$ has density function

   $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

   The moment generating function of X is

   $$M(t) = e^{\mu t + \sigma^2 t^2/2}.$$

   Prove the following, using the moment generating function.

   - $E[X] = \mu$.
   - $\text{Var}[X] = \sigma^2$.
   - If $X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma^2)$, and X_1 and X_2 are independent, then $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

11. 7-2

    Suppose that $x_1, x_2, \cdots x_n$ are a random sample from a uniform distribution Unif$[0, \beta]$ with unknown parameter β. Show that the maximum likelihood estimator of β is $\max x_i$ and that this estimator can be made unbiased by multiplying it by $(n+1)/n$.

12. 8-1 Define the gamma function for $r > 0$ by $\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx$. Prove that

    - $\Gamma(r+1) = r\Gamma(r)$ if $r > 0$.
    - For integer $n > 0, \Gamma(n+1) = n!$
    - $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Group 3

13. 8-4 Let X_1, X_2 be a random sample from $N(0,1)$, with sample mean $\overline{X}$ and sample variance S^2. Prove the following:

    - $X_1^2 + X_2^2 = 2\overline{X}^2 + S^2$.
    - $E[\overline{X}^2 S^2] = E[\overline{X}^2]E[S^2]$
      (Use the fact that if X is standard normal, $E[X^2] = 1$ and $E[X^4] = 3$.

14. 8-5 Let $X_1, X_2, \cdots X_n$ be a random sample from $N(\mu, \sigma^2)$, with sample mean $\overline{X}$ and sample variance S^2. It continues to be true that $\overline{X}$ and S^2 are independent random variables. Define

    $$U = \frac{1}{\sigma^2}\sum_{i=1}^{k}(X_i - \mu)^2; V = \frac{1}{\sigma^2}\sum_{i=1}^{k}(X_i - \overline{X})^2; W = \frac{1}{\sigma^2}n(\overline{X} - \mu)^2.$$

    Prove the following:

    - $U = V + W$.
    - $(n-1)S^2/\sigma^2$ has a chi-square distribution with $n-1$ degrees of freedom.

15. 8-7 Prove that if $X_1, X_2, \cdots X_n$ are a random sample from $N(\mu, \sigma^2)$, then

    $$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

    has a t distribution with $n-1$ degrees of freedom.

16. 8-8 Student t confidence interval

    Let $X_1, X_2, \cdots X_n$ be a random sample from $N(\mu, \sigma^2)$, with both μ and σ unknown. The sample mean is $\overline{X}$; the sample variance is S^2.

    Let q denote the $(1-\alpha/2)$ quantile of the Student t distribution with $n-1$ degrees of freedom. By symmetry, $-q$ is the $\alpha/2$ quantile.

    Define random variables

    $$L = \overline{X} - \frac{qS}{\sqrt{n}}; U = \overline{X} + \frac{qS}{\sqrt{n}}.$$

    Prove that $P(L > \mu) = \alpha/2$ and $P(U < \mu) = \alpha/2$,
    so that $P(L \leq \mu \leq U) = 1 - \alpha$. and $[L, U]$ is a $1-\alpha$ confidence interval.

17. 9-1 Covariance and correlation

    The covariance of random variables X and Y is defined as
    $\mathrm{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)]$

    The correlation coefficient of random variables X and Y is defined as

    $$\rho(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}.$$

    (a) Prove that $\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y]$.

    (b) Prove that $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X,Y)$.

    (c) Prove that $|\rho(X,Y)| \leq 1$.

18. 9-2 Least-squares regression

    You have values x_i of a "predictor" and matching values y_i of a "response." Your goal is to minimize the sum of squares of the prediction errors,

    $$g(a,b) = \sum_{i=1}^{n} (a + bx_i - y_i)^2.$$

    Prove that

    $$b = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2}, a = \overline{y} - b\overline{x}.$$

Group 4

19. 10-1

    (a) State and prove Bayes' theorem for two events A and B, neither of which has probability zero.

    (b) Specialize to the case where X are some data (one or more random variables) from a discrete probability distribution that is specified by one or more parameters collectively represented by θ. The event A is "the parameter value is θ_j"; the event B is "the observed data are X."

20. 10-2

    Suppose that random variable X has a beta distribution, i.e. its probability density function is

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, 0 \leq x \leq 1$$

Knowing that $\displaystyle\int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx = 1,$ prove that

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

21. 10-3

    Suppose that the parameter θ for a binomial distribution has the prior distribution $\theta \sim$ Beta(α, β) and the data are given by the binomial distribution $X \sim$ Binom(n, θ). Prove that the posterior distribution $\theta | X \sim$ Beta$(\alpha + x, \beta + n - x)$.

$$\text{Reminder: } \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} d\theta = 1.$$

22. 11-1

Pivotal statistic – location parameter

Suppose that $X_1, X_2, \cdots X_n$ are random variables from a distribution with parameter θ. A *pivotal statistic* is a function $h(X_1, X_2, \cdots X_n, \theta)$ whose distribution does not depend on θ or on any parameters with unknown values. The parameter θ is called a *location parameter* if the distribution of $X - \theta$ does not depend on θ.

(a) Show that for a normal distribution with known σ, μ is a location parameter and that given one sample X from this distribution, there are random variables L and U, with values that depend on X but not on θ, such that $P(\mu < L) = 0.025$ and $P(\mu > U) = 0.025$

(b) Show that if $X \sim \text{Unif}(\theta - 1, \theta + 1)$, then θ is a location parameter, and explain how to determine L and R for a 90% confidence interval.

23. 11-2

Pivotal statistic – scale parameter

A parameter θ is called a *scale parameter* if the distribution of X/θ does not depend on θ.

(a) Show that if X has the distribution $N(2\sigma, \sigma^2)$, then σ is a scale parameter, and find a formula for L if you want a 97.5% one-sided confidence interval for σ.

(b) Show that if X has the distribution $\text{Unif}(0, \theta)$, then θ is a scale parameter. Find formulas for L and U that can be used if you know only $X = \max(X_1, X_2)$ and want a 92% confidence interval.

24. 12-1

A modern-day Robinson Crusoe has been shipwrecked on an island which may lie off the coast of Brazil. Among the few items that he was able to salvage from his sinking ship are two books: one containing tables of the distribution functions for the standard normal distribution and for Student t distributions with various numbers of degrees of freedom and one named Wild Plants of the Atlantic Islands.

Like the original Crusoe, he makes raisins from the local wild grapes. From a random sample of n vines, he gets an average yield of $\overline{x}$ grams of raisins. From his botany book, he learns that on islands off the coast of Brazil, the yield of raisins has a normal distribution with mean μ and variance σ^2.

(a) How, given available resources, can he test the null hypothesis that his island is off the coast of Brazil?

(b) Suppose that the botany book includes only the value of μ, not of σ^2, but that modern Crusoe has calculated the sample standard deviation s of his raisin yields. How does he test the null hypothesis that his island is off the coast of Brazil?