Theory Comprehensive Exams

Nick Sun, Ying Dai

April 5, 2020

1 Introduction

This is planned to be a running list of comprehensive exam questions from Oregon State University written out in a single LaTeX document. My eventual goal is to have a document that is easily GREP-able for questions of a certain type, for example, questions that required using Jensen's Inequality or an integral that had to be evaluated using the Beta kernel. I also want to get solutions and hints are written up as well.

Contents

2 Exam 2000

2.1 Problem 1

N is a Poisson random variable where λ is chosen from an Exp(1) distribution

Part a. Find the unconditional probabilities $P[N = x]$ for $x = 0, 1, 2, \ldots$.

$$P[N = x] = \int_0^\infty P(X|\lambda)P(\lambda)d\lambda$$

HINT: Integral makes use of Gamma kernel trick $\int_0^\infty exp(-2\lambda)\lambda^x = \Gamma(x + 1)(\frac{1}{2})^{x+1}$

Part b. Find the unconditional expected value of N

$$E[N] = \sum_{x=0}^{\infty} xP(N = x) = \sum_{x=0}^{\infty} x(\frac{1}{2})^{x+1}$$

HINT: Makes use of $2E[N] - E[N]$ and the series identity $\sum_{x=1}^{\infty}(\frac{1}{2})^x = 1$

Part c. Is the unconditional variance of N less than 1, equal to 1, or greater than 1?

HINT: Makes use of the law of total expectation and total variance

$$E[N] = E[E[N|\lambda]] = E[\lambda] = 1$$

and

$$Var[N] = E[Var[N|\lambda]] + Var[E[N|\lambda]] = E[\lambda] + Var[\lambda] = 1 + 1 = 2$$

2.2 Problem 2

Suppose that $X_1, \ldots, X_n$ are i.i.d. with a shifted exponential distribution

Part a. Find a two dimensional sufficient statistic for (α, β).

HINT: Using the factorization theorem, we can split the likelihood function into separate functions for α and β. We get $T = (X_{(1)}, \sum_{i=1}^n x_i)$ as our two dimensional sufficient statistic.

Part b. Find the MLEs for α and β.

HINT: $f(x|\alpha, \beta)$ increases as α increases, so the MLE of α is $X_{(1)}$. After doing some calculus, the MLE for β is $\frac{1}{n}\sum(x_i - X_{(1)})$

Part c. Find the minimal sufficient statistic for β

Using Lehmann-Scheffe (1), we find that $\sum x_i$ or $\bar{X}$ is the minimal sufficient statistic for β.

Part d. Find the UMVUE of β

HINT: Since $\sum x_i$ is the minimal sufficient statistic for β and 1-1 function $\sum(x_i - \alpha) \sim Gamma(n, \beta)$ is our complete sufficient statistic, let $T = \sum x_i - \alpha$ where $E[T] = n\beta$, so therefore $W = \frac{1}{n}\sum x_i - \alpha$ is he UMVUE for β.

Part e. Find the UMVUE for β^2

HINT: Try $E[\sum(x_i-\alpha)^2] = \sum(Var[x_i-\alpha]+E[x_i-\alpha]^2) = \sum \beta^2+\beta^2 = 2n\beta^2$. Therefore $W = \frac{1}{2n}\sum(x_i - \alpha)^2$ is the UMVUE for β^2.

2.3 Problem 3

TOPIC: Probability Modelling

A component has an exponential lifetime with an expected life of 10 years. A unit consists of four identical, independent components.

Part a. A unit has 4 components wired in *series* if the failure of any single component makes the unit fail. What is the distribution of the units lifetime?

HINT: Use cumulative distribution function of $Y = min(X_1, \ldots, X_4)$.

$$1 - P(X_1 > y)P(X_2 > y)P(X_3 > y)P(X_4 > y) = 1 - exp(\frac{y}{10})^4$$

So $Y \sim Exp(\frac{5}{2})$

Part b. What is the units expected lifetime if the unit has its components in series?

Part c. A unit has the four components in *parallel* if the unit fails only when all the components fail. What is the distribution of the units lifetime if the unit has its components in parallel?

HINT: Let $W = max(X_1, \ldots, X_4)$. Using cumulative distribution functions, we find that $P(W < w) = P(X_1 < w)P(X_2 < w)P(X_3 < w)P(X_4 < w) = (1 - exp(-\frac{w}{10}))^4$ and $f_w(w) = \frac{2}{5}(1 - exp(-\frac{w}{10}))^3 exp(-\frac{w}{10})$.

Part d. What is the unit's expected lifetime if the unit has its components in parallel?

HINT: Should be $\frac{2}{5}(10^2 - 3(5^2) + 3(\frac{10}{3})^2 - (\frac{10}{4})^2) = \frac{125}{6}$

2.4 Problem 4

TOPIC: UMPs with the $f(x|\theta) = \theta x^{\theta-1}$ which is the Beta(θ, 1) distribution.

Let $X_1, \ldots, X_n$ is i.i.d. Beta(θ, 1) for $0 < x < 1$ for some θ ¿ 0. Let $W = -ln(\prod x_i)$.

It can be shown that $2\theta W \sim \chi^2_{2n}$

Part a. Find the UMP size α test for testing $H_0 : \theta = 1$ vs. $H_1 : \theta > 1$

HINT: The monotone likelihood ratio is decreasing w.r.t. $\prod x_i$ so we reject for large values of $\prod x_i$ or $-ln(\prod x_i) < k^*$

Part b. Find the MLE of θ.

HINT: The MLE of θ is $-\frac{n}{ln(\prod x_i)}$

Part c. Find the Likelihood Ratio Test for testing $H_0 : \theta = 1$ vs $H_1 : \theta \neq 1$. Show that it rejects if W is too large or too small.

HINT: Find the Likelihood function with the MLE of θ plugged in and then take the derivative. It will be seen that we should reject H_0 if W is too large and too small compared to the sample size n.

Part d. Find a $1 - \alpha$ CI based on the LRT in part c. Find constraints which determine the lower bound L(W) and upper bound U(W) of the confidence interval.

2.5 Problem 5

TOPIC: Showing $\bar{X}$ and s^2 is independent.

Suppose that $X_1, \ldots, X_n$ is a random sample of n independent observations from a $N(\mu, \sigma^2)$ distribution where μ and σ^2 are unknown parameters.

Part a. What is the distribution of $\bar{X}$?

Should be $N(\mu, \frac{\sigma^2}{n})$

Part b. What is the distribution of s^2?

Should be $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

Part c. Is the correlation between $\bar{X}$ and s positive, negative, or 0?

HINT: Use Basu's theorem. $\bar{X}$ is a CSS for μ and s^2 is ancillary for μ. The estimates are independent and the correlation is 0.

Part d. Suppose that $Y_1, \ldots, Y_m$ is a random sample of m independent observations from a $N(0, \sigma^2)$ distribution where σ^2 is the same unknown parameter as above. Using only the Ys, find an unbiased estimator of σ^2 with variance $\frac{2\sigma^4}{m}$.

HINT: Using $\sum_{i=1}^{m}(\frac{Y_i}{\sigma})^2 \sim \chi^2_m$ which gives us $E[\frac{1}{m}\sum_{i=1}^{m} Y_i^2] = \sigma^2$

e. Let μ_0 be hypothesized value for μ. Suppose the sample of Xs and the sample of Ys are independent from one another. Using both samples, construct a test statistic T whose distribution when $\mu = \mu_0$ is a t-distribution with n+m-1 df.

HINT: Use $\frac{1}{n}\sum \frac{x_i}{\sigma} \sim N(\mu, \frac{1}{n})$ and $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$ and $\sum_{i=1}^{m}(\frac{Y_i}{\sigma})^2 \sim \chi^2_m$.

We get $T = \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sqrt{(n-1)s^2+\sum Y_i^2/(n+m-1)}} \sim t_{n+m-1}$ and note that the numerator is independent of the denominator.

3 Exam 2001

3.1 Problem 1

TOPIC: Properties of the Standard Normal distribution

Let $X_1, \ldots, X_n$ be i.i.d. N(0, 1)

Part a. Find the correlation between X_1 and $\bar{X}$

HINT: $cov(x_1, \bar{X}) = \frac{1}{n}cov(x_1, x_1) = \frac{1}{n}$ so correlation between X_1 and $\bar{X}$ is $\frac{\frac{1}{n}}{\sqrt{1}\sqrt{\frac{1}{n}}} = \frac{1}{\sqrt{n}}$

Part b. Find the distribution of $\bar{X}$ given $X_1 = x_1$

HINT: Use the fact that $\sum_{i=2}^{n} x_i \sim N(0, n-1)$ and set that pdf for the value $n\bar{X} - x_1$

Part c. Find the distribution of X_1 given that $\bar{X} = \bar{x}$

HINT: Use that fact that X_1 is independent of $X_2, \ldots, X_n$ on $f(X_1 = x_1 | \sum_{i=2}^{n} x_i = n\bar{X} - X_1)$ to get the pdf

3.2 Problem 2

TOPIC: Properties of the Poisson distribution

$X_1, \ldots, X_n$ is i.i.d. Poisson(λ)

Part a. Find a minimal sufficient statistic

HINT: Using Lehmann-Scheffe (1), we find that $T = \sum X_i$ is our minimal sufficient statistic

Part b. Let $\theta = P_\theta(X_1 = 0)$. Find an unbiased estimator of θ that is a function on the minimal sufficient statistic.

HINT: Use $E[I(X_1 = 0)|Y = y] = P(X_1 = 0| \sum X_i = y) = (\frac{n-1}{n})^y$. Then we plug in $\sum X_i$ for the value of y

Part c. Find the variance of the estimator in part (b).

HINT: Find $E[W^2]$ and $E[W]$ which requires multiplying a constant into a summation to get a Poisson pmf to get $E[W^2] = exp(-2\lambda + \frac{\lambda}{n})$

Part d. Let Y denote the number of observations X_i that are 0. Find a function of Y that is unbiased for θ.

HINT: $Y \sim Binomial(n, p = \theta)$, so $\frac{Y}{n}$ is unbiased for θ

Part e. Find the variance of the estimator in part d.
HINT: $Var[\frac{Y}{n}] = \frac{\theta(1-\theta)}{n} = \frac{exp(-\lambda)(1-exp(-\lambda))}{n}$

Part f. Which of the two estimators in parts (b) and (d) has the smaller variance?
HINT: (b) is the UMVUE so yeah...

3.3 Problem 3

TOPIC: UMP test for Beta($\theta+1$, θ)

Part a. Find UMP test at level α for $H_0 : \theta = \theta_0$ vs. $H_1 : \theta > \theta_0$
HINT: The ratio should be monotone decreasing with respect to $\prod X_i$. To get the critical value, use the fact that $Y = -ln(x) \sim exp((\theta+1)^{-1})$ so $\sum Y_i \sim Gamma(n, (\theta+1)^{-1})$

Part b. Find the rejection region of the UMP test when $\alpha = .05$ and $\theta_0 = 2$, $n = 10$.

Part c. Find a 95% confidence interval for θ when $n = 10$ by inverting the UMP test in part a.
HINT: Use the fact that $-(\theta+1)ln(\prod X_i) \sim Gamma(n, 1)$

3.4 Problem 4

TOPIC: Finding the MLE and LRT of the exponential($\frac{1}{\beta}$), $x > 0, \beta > 0$

Part a. Find the MLE for β
HINT: The MLE of β is $\frac{n}{\sum X_i}$

Part b. Derive the LRT for testing $H_0 : \beta = \beta_0$ $H_1 : \beta \neq \beta_0$
HINT: IF $\beta_0 > \beta_{MLE}$ then $\Lambda(X)$ is monotonically decreasing w.r.t. $\sum X_i$. The reverse if $\beta_0 < \beta_{MLE}$.

Part c. What is the asymptotic distribution of -2ln($\Lambda(X)$)
It is known to be χ^2_1

Part d. If $n = 50$, $\beta_0 = 1$, $\sum X_i = 65$, $\alpha = .05$, do we acept or rejet the H_0 from (b)?
We should fail to reject H_0.

Part e. If $n = 50$, $\beta_0 = 1$, $\sum X_i = 65$, $\alpha = .05$, do we acept or rejet the H_0 using the LRT asymptotic test in part (c)?
We should fail to reject H_0.

4 Exam 2002

4.1 Problem 1

TOPIC: Finding the UMP test for the Rayleigh distribution
$X_1, \ldots, X_n \sim f(x|\theta) = 2\theta^{-1} xexp(-\frac{x^2}{\theta})$

Part a. Derive a UMP test for $H_0 : \theta \geq \theta_0$ vs $H_1 : \theta < \theta_0$
HINT: Should find that the ratio is monotone increasing function w.r.t. $\sum X_i^2$

Part b. Derive the likelihood ratio test for $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$
HINT: The unconstrained MLE comes out to be $\frac{\sum X_i^2}{n}$

Part c. Derive the likelihood ratio test for $H_0 : \theta > \theta_0$ vs. $H_1 : \theta < \theta_0$
HINT: Note that the constrained MLEs has two cases: if $\theta_0 \leq \theta_{MLE}$, then $\theta_{0,MLE} = \theta_{MLE}$. IF $\theta_0 > \theta_{MLE}$, then $\theta_{0,MLE} = \theta_0$

Part d. Are the tests in (a) and (c) equivalent? Justify your answer.
HINT: Take the derivative of the likelihood ratio test in part (c) and it can be seen that $\Lambda(X)$ is the monotone increasing function w.r.t. $\sum X_i^2$ so the tests are equivalent.

4.2 Problem 2

TOPIC: Properties of Normal Distributions with different location parameters
$Y_1, \ldots, Y_n$ is i.i.d. $N(\beta x_i, \sigma^2)$ with known constants $x_1, \ldots, x_n$.

Part a. Find a two dimensional sufficient statistic for (β, σ^2).
HINT: We should find that $(\sum y_i^2, \sum x_i y_i)$ is the two-dimensional sufficient statistic.

Part b. Find the maximum likelihood estimates for β and σ^2
HINT: After doing a few partial derivatives, we get that $\beta_{MLE} = \frac{\sum x_i y_i}{x_i^2}$ and $\sigma^2_{MLE} = \frac{1}{n}(\sum y_j - \beta_{MLE} x_j)^2$

Part c. Find the UMVUE for β
HINT: After doing some factorization of the PDF, we see that the CSS for β is $\sum x_i y_i$. Taking the expectation of this and then correcting it, we find that $W = \frac{\sum x_i y_i}{\sum x_i^2}$ is the UMVUE for β

Part d. Find the distribution of the MLE for β
HINT: $x_i y_i \sim N(\beta x_i^2, x_i^2 \sigma^2)$ so $\frac{\sum x_i y_i}{\sum x_i^2} \sim N(\beta, \frac{\sigma^2}{\sum X_i^2})$

Part e. How would you obtain a $100(1-\alpha)$

HINT: To begin, we need a pivotal quantity.

Try $T = \frac{\frac{\sum x_i y_i}{\sum x_i^2} - \beta / \sqrt{\frac{\sigma^2}{\sum x_i^2}}}{\frac{(n-1)s^2}{\sigma^2}/(n-1)} \sim t_{n-1}$

4.3 Problem 3

TOPIC: Properties of a random sample

$X_1, \dots, X_n$ is a random sample.

Part a. Suppose the only thing you known about the density of the Xs is that $Var(\frac{1}{\sqrt{X_i}}) < \infty$. Suggest a consistent estimator of $E[\frac{1}{\sqrt{X_i}}]$

HINT: Using WLLN, $\frac{1}{n}\sum \frac{1}{\sqrt{x_i}} \rightarrow_p E[\frac{1}{\sqrt{X_i}}]$ since the variance is finite, so the sample mean is our consistent estimator.

Part b. Suppose the denisty of X_i is $f_\theta(x) = \frac{1}{\theta} exp(-\frac{x}{\theta}), x > 0$. Estimate $E[\frac{1}{\sqrt{X_i}}]$ using calculus and come up with an MLE.

HINT: The MLE will be the sample mean. The expected value an ve found using the Gamma kernel trick: $\int_0^\infty \frac{1}{\sqrt{X_i}} \frac{1}{\theta} exp(-\frac{x}{\theta}) = \Gamma(\frac{1}{2})\theta^{-\frac{1}{2}} = \sqrt{\frac{\pi}{\theta}}$

Part c. Are the estimates in (a) and (b) the same?

4.4 Problem 4

This question was ridiculous and I didn't do it (see "civil disobedience").

4.5 Problem 5

TOPIC: Some results on limiting distributions

Part a. Let X_n be a χ^2_n r.v. Show that the limiting distribution of $\sqrt{n}(\frac{X_n}{n} - 1)$ is N(0, 2) as $n \rightarrow \infty$.

HINT: Let $X_n = \sum_{i=1}^n Z_i^2$ where $Z_i \sim N(0, 1)$. $E[Z_i^2] = 1$ and $Var[Z_i^2] = 2$. From CLT, then $\sqrt{n}(\frac{X_n}{n} - 1) \rightarrow_d N(0, 2)$

Part b. Let $Y_1, \dots, Y_n$ be i.i.d. $N(\mu, \sigma^2)$ and let s_n^2 be the sample mean. Use the result in (a) to derive the limiting distribution of $\sqrt{n}(s_n^2 - \sigma^2)$ as $n \rightarrow \infty$.

HINT: Use the fact that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} = \sum_{i=1}^{n-1} W_i$ and then invoke CLT to get the limiting distribution.

5 Exam 2003

5.1 Problem 1

TOPIC: MLE of the Pareto Distribution

$X_1, \ldots, X_n$ are i.i.d. $f(x) = \frac{aK^a}{x^{a+1}}, x \geq K$ and a, K are positive

Part a. Find a pair of jointly sufficient statistics for a and K

HINT: Factoring the pdf apart gives us $T = (\prod X_i, X_{(1)})$ as our jointly sufficient statistic.

Part b. Find he MLEs for a and K

HINT: Inspection shows that the likelihood increases as K increases, so $K_{MLE} = X_{(1)}$. Differentiating the log likelihood shows that $a_{MLE} = \frac{n}{\sum ln(\frac{x_i}{k})}$

Part c. Assuming K is known to be 1, find the MLE for $A = \frac{1}{a}$

HINT: Since MLEs are invariant, $A_{MLE} = \frac{\sum ln(x_i)}{n}$.

Part d. Find the mean and variance of A.

HINT: The distribution $Y_i = ln(X_i) \sim exp(\frac{1}{a})$ and $\sum Y_i \sim Gamma(n, \frac{1}{a})$. This gives us that $E[A_{MLE}] = \frac{1}{a}$ and $Var[A_{MLE}] = \frac{1}{a^2 n}$

Part e. Compute the information $I_{x_i}(a)$ and find the CRLB for the variance of the unbiased estimators of A.

HINT: Recall that the pdf is given as $f(x|a) = \frac{a}{x^{a+1}}, x \geq 1$. Then $I_{x_i}(a) = \frac{1}{a^2}$ and the CRLB $= \frac{[g'(a)]^2}{nI_{x_i}(a)} = \frac{1}{a^4}\frac{a^2}{n} = \frac{1}{na^2}$

Part f. Is A_{MLE} the UMVUE of $\frac{1}{a}$?

HINT: Yes, since it achieves the CRLB. Alternatively, you can factor $f_y(y) = aexp(-ay)$ apart to see the CSS.

Part g. What is the asymptotic distribution of A_{MLE}?

HINT: Use CLT. Since $A_{MLE} \sim \Gamma(\frac{1}{a}, \frac{1}{na^2})$, then $\sqrt{n}(A_{MLE} - \frac{1}{a}) \sim N(0, \frac{1}{a^2})$

5.2 Problem 2

TOPIC: Joint and marginal distributions of Exponential r.v.s

Let X_1, X_2, X_3 be i.i.d. with $f(x) = \frac{1}{\theta}exp(-\frac{1}{\theta}), x > 0, \theta > 0$

Furthermore, define $Y_1 = X_1, Y_2 = X_1 + X_2, Y_3 = X_1 + X_2 + X_3$

Part a. Determine the joint distribution of Y_1, Y_2, Y_3.

HINT: This requires calculating the Jacobian which turns out to be
$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$
the determinant of which is 1 since we can row reduce this to an identity matrix.

Afterwards, we can just plug in to the original joint pdf (easy since they are all independent) with $X_1 = Y_1, X_2 = Y_2 - Y_1, X_3 = Y_3 - Y_2$. The final answer should be $\frac{1}{\theta^3} exp(-\frac{y_3}{\theta}), y_3 \geq y_2 \geq y_1 \geq 0$

Part b. Determine the marginal distribution of Y_3

HINT: This just involves integrating the joint pdf over dy_1 (limits 0 to y_2) and dy_2 (limits 0 to y_3). The final answer should be $\frac{y_3^2}{2\theta^3} exp(-\frac{y_3}{\theta})$

Part c. Determine the conditional distribution of Y_1 and Y_2 given $Y_3 = y_3$.

HINT: This just involves employing the definition of conditional distributions so $f_{y_1,y_2,y_3}(y_1, y_2|y_3) = \frac{f_{y_1,y_2,y_3}(y_1,y_2,y_3)}{f_{y_3}(y_3)} = \frac{2}{y_3^2}, y_3 > y_2 > y_1 > 0$

Part d. Determine the conditional expectation $E[Y_2|Y_3 = y_3]$

HINT: First we need to get $f_{Y_2|Y_3}(y_2|y_3) = \int_0^{y_2} \frac{2}{y_3^2} dy_1 = \frac{2y_2}{y_3^2}, y_3 > y_2 > 0$ and then we calculate the conditional expectation using the new conditional distribution $E[Y_2|Y_3 = y_3] = \int_0^{y_3} y_2 \frac{2y_2}{y_3^2} dy_2$

Part e. What does (a), (b), (c) tell you about how you should use these variables to estimate θ?

The only value you need is y_3! It is the complete sufficient statistic!

5.3 Problem 3

TOPIC: UMP test for a Rayleigh distribution

$X_1, \ldots, X_n$ be i.i.d. $f(x|\theta) = 2\theta^{-1} xexp(-\frac{x^2}{\theta}) I_{x>0}, \theta > 0$

Part a. Find the UMP level α test for $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta \geq \theta_0$

HINT: Similar to previous years. Should find that this is a decreasing function of $\sum X_i^2 \sim \Gamma(n, \theta)$

Part b. Find an appropriate pivot quantity and use it to construct a two-sided equal tail $(1 - \alpha)100\%$ CI for θ.

HINT: An appropriate pivot quantity is $\frac{\sum X_i^2}{\theta} \sim \Gamma(n, 1)$

Part c. Find the maximum likelihood estimator of θ

HINT: A little differentiation should show that $\theta_{MLE} = \frac{\sum X_i^2}{n}$

Part d. Find the likelihood ratio test of asymptotic size α for testing $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$

HINT: Should get $\Lambda(X) = \left(\frac{\sum X_i^2}{n\theta_0}\right)^n exp(-\frac{\sum X_i^2}{\theta_0} + n)$. Then take the log of this and multiply by -2 to get your χ^2 statistic.

5.4 Problem 4

TOPIC: Convergence with Normal Distribution

$X_1, \ldots, X_n$ are i.i.d. $N(\mu, 1)$

Part a. Find the asymptotic distribution of the sample mean $\bar{X}$ as $n \to \infty$

HINT: From CLT, $\sqrt{n}(\bar{X} - \mu) \to_d N(0, 1)$

Part b. Use part (a) to perform a large sample test of $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

HINT: They are just asking for a z-test here. Use $Z = \frac{\bar{X}-\mu}{\sqrt{\frac{1}{n}}} \sim N(0, 1)$ under H_0

Part c. Does $\bar{X}^2 exp(-\bar{X})$ converge to a limit in probability?

HINT: By continuous mapping theorem (Mann-Wald).

Part d. Find a large sample approximation for the distribution of $\bar{X} exp(-\bar{X})$.

HINT: Use the Delta method to get distributional convergence.

5.5 Problem 5

TOPIC: Moment generating functions and Poisson r.v.s

Let A, B, C be independent Poisson r.v.s with mean α, β, γ respectively.

Part a. Use MGFs to determine the distribution of X = A + C

HINT: $M_X(t) = exp(\alpha(e^t - 1))exp(\gamma(e^t - 1)) = exp((\alpha + \gamma)(e^t - 1))$ so $X \sim$ Poisson$(\alpha + \gamma)$

Part b. Let X = A + C, Y = B + C. Find an expression of the correlation between X and Y in terms of the parameters α, β, γ

HINT: cor(X, Y) = cor(A + C, B + C) = cor(A, B) + cor(A, C) + cor(C, B) + var(C) = var(C) = γ

Part c. Let $\lambda = E[X], \mu = E[Y]$. Suppose that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are a random sample of pairs of counts following the model above. The difference in sample means $\bar{X} - \bar{Y}$ is an unbiased estimator of $\lambda - \mu$. Find an unbiased estimator of Var$(\bar{X} - \bar{Y})$

HINT: Var$(\bar{X} - \bar{Y})$ = Var$\bar{X}$ + Var$\bar{Y}$ - 2cov$(\bar{X}, \bar{Y}) = \frac{\alpha+\gamma}{n} + \frac{\beta+\gamma}{n} - 2(\frac{\gamma}{n}) = \frac{\alpha+\beta}{n}$

6 Exam 2004

6.1 Problem 1

TOPIC: MLEs of a Shifted Exponential distribution

$X_1, \ldots, X_n$ be i.i.d. $f(x|\theta) = exp(-(x-\theta)), \theta \leq x < \infty$

Part a. Write the likelihood function and find the MLE of θ

HINT: The MLE is $X_{(1)}$

Part b. Find the distribution function and the pdf of θ_{MLE}

HINT: Using $P(X_{(1)} > x) = \prod P(X_i > x) = exp(-n(x-\theta))$ we find that $F_{X_{(1)}}(x) = 1 - exp(-n(x-\theta))$ so $f_{X(1)}(x) = nexp(-n(x-\theta)), \theta \leq x < \infty$ or alternatively can use the formula for the pdf of the minimum order statistic.

Part c. Find $E[\theta_{MLE}]$. Is θ_{MLE} unbiased for θ

HINT: This is easily done since $X_{(1)}$ belongs to a location scale family so $n(X_1 - \theta) \sim exp(1)$ giving us $E[X_{(1)}] = \frac{1}{n} + \theta$

Part d. Find a $100(1-\alpha)\%$ CI for θ.

HINT: Use $n(X_{(1)} - \theta) \sim Exp(1)$ as your pivot!

6.2 Problem 2

TOPIC: Moment generating functions with Chi-squared distribution

Let $X \sim \chi^2_p$ with $f(x|p) = \frac{a}{\Gamma(p/2)2^{p/2}} x^{p/2-1} exp(-\frac{x}{2})$ and let $Y \sim \chi^2_q$ and independent of X.

Part a. Find the MGF of X

HINT: This involves using the χ^2 kernel.

$M_X(t) = E[e^{tX}] = \int_0^\infty e^{tX} \frac{1}{\Gamma(p/2)2^{p/2}} x^{p/2-1} exp(-\frac{x}{2})dx$

$\frac{\Gamma(p/2)(\frac{1}{2}-t)^{-p/2}}{\Gamma p/2 2^{p/2}} \int_0^\infty \frac{1}{\Gamma(p/2)(1/2-t)^{-p/2}} x^{p/2-1} exp(-(1/2-t)x)dx$

$\frac{1}{(1-2t)^{p/2}}, t < \frac{1}{2}$

Part b. Show that $X + Y \sim \chi^2_{p+q}$

HINT: This is shown easily by $M_{X+Y}(t) = M_X(t)M_Y(t)$

Part c. Find the density of $\frac{X}{Y}$

HINT: Using the fact that $\frac{X/p}{Y/q} \sim F_{p,q}$, let W $= \frac{X}{Y} = \frac{p}{q} F_{p,q} \rightarrow \frac{q}{p} W = F_{p,q}$ Then plugging this in to the pdf of an F distribution which is ugly to begin with, we get some really ugly looking pdf.

Part d. Does $\frac{X}{p}$ have a limit as $p \to \infty$?

HINT: Sure. Since X is just a sum of p χ_1^2 r.v.s, we can use WLLN to show that $\frac{X}{p} \to_p 1$

6.3 Problem 3

TOPIC: UMP tests with Poisson r.v.s

X and Y are independent r.v.s where $X \sim Poisson(\lambda)$ and $Y \sim Poisson(\lambda + 1)$

Part a. Find the test statistic and the most powerful test of$H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1$. Give the form of the rejection region.

HINT: $T = \left(\frac{\lambda_0}{\lambda_1}\right)^x \left(\frac{\lambda_0+1}{\lambda_1+1}\right)^Y$ and we reject when $T < k$

Part b. Show that the test in part (a) can be expressed as "Reject H_0 iff aX + bY ¿ c".

HINT: If you take the log, you find that $ln(T) = Xln(\frac{\lambda_0}{\lambda_1}) + Yln(\frac{\lambda_0+1}{\lambda_1+1})$

Part c. Is there a uniformly most powerful test of $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda \neq \lambda_0$?

HINT: No, cannot have a single test which has both rejection regions.

Part d. If λ is large, X has an approximately normal distribution. Why?

HINT: Use MGFs. $E[exp(\frac{xt}{\lambda})]$ is equal to the normal MGF since the infinite summation is equal to 1. Therefore if λ is large, then $\frac{X}{\lambda} \sim N(1, \frac{1}{\lambda}) \to X \sim N(\lambda, \lambda)$

Part e. Let $\lambda_0 = 20$ and $\lambda_1 = 30$ and are relatively large. We observe X = 22 and Y = 27. Calculate the p-value of the test from part (a) in the orm of of $\phi(d)$

f. Suppose we observe X = 22 and Y = 27. The MLE of λ is 23.96. Consider the LRT of $H_0 : \lambda = 20$ vs. $H_1 : \lambda \neq 20$. Express the approximate p-vlue of this test in the form of 1 - F(d) where F is the cdf of the χ^2 distribution.

6.4 Problem 4

TOPIC: MLEs, Least Square Estimators, UMVUEs of Poisson r.v.s with different constants

$X_1, \ldots, X_n$ be i.i.d. Poisson with $E[X_i] = \lambda z_i$ where z_i are known constants and $\lambda > 0$ is an unknown parameter.

Part a. Find the least squares estimate of λ. Is it unbiased?

HINT: Taking the derivative of $\sum(X_i - E[X_i])^2 = \sum X_i^2 - 2\lambda \sum X_i z_i + \lambda^2 \sum z_i^2$ gives us $\hat{\lambda} = \frac{\sum z_i x_i}{\sum z_i^2}$. Taking the expectation of this shows it is unbiased.

Part b. Is $\hat{\lambda}$ the UMVUE estimator for λ?

HINT: Since the Poisson pdf belongs to an exponential family, the CSS for λ can be shown to be $\sum X_i$ (note that our least squares estimator is *not* a function of this statistic). Taking the expectation of this and correcting it, we can see that $T = \frac{\sum X_i}{\sum z_i}$ is the UMVUE for λ

Part c. Show that the UMVUE achieves the CRLB for unbiased estimators of λ.

HINT: We can calulate that the information on λ contained in the sample is $\frac{\sum z_i}{\lambda}$. Therefore the CRLB is $\frac{\lambda}{\sum z_i}$. The variance of $\frac{\sum X_i}{\sum z_i}$ is equal to this, so the CRLB is met.

Part d. Suppose that $z_i = i, i = 1, \ldots, n$. Is $\hat{\lambda}$ a consistent estimator?

HINT: We are given that $\sum_{i=1}^{n} \frac{(n+1)n}{2}$ so the variance of $\hat{\lambda}$ is $\frac{2\lambda}{(n+1)n}$ which goes to 0 as $n \to \infty$. Therefore $\hat{\lambda}$ is consistent.

6.5 Problem 5

TOPIC: P-values

We want to test H_0 vs H_1. We are able to find the results of 5 studies using the test statistics $T_1, \ldots, T_5$ Assume large T_i indicated support for H_1. Under H_0, each T_i has a cdf F_i.

Part a. SHow that the p-value for the ith test P_i has a U(0, 1) distribution under H_0.

HINT: $P_i = 1 - F_i(T_i) = P_{H_0}(1 - F_i(T_i) < \alpha) = P_{H_0}(T_i > F_i^{-1}(1-\alpha)) = \alpha$, therefore P_i has a U(0, 1) distribution.

Part b. Find the distribution of $-\sum_{i=1}^{5} ln(P_i)$ under H_0.

HINT: Let $Y_i = -ln(P_i) \sim exp(1)$ so $\sum_{i=1}^{5} Y_i \sim \Gamma(5, 1)$

Part c. How might $-\sum_{i=1}^{5} ln(P_i)$ be used to test H_0 vs H_1?

HINT: Use $\Gamma(5, 1)$ as a reference distribution for the test statistc $T = -ln(\sum ln(p_i))$

Part d. Suppose that the researcher is able to locate the results of n studies and calculate the p-value from each study as above. Suppose $n \to \infty$. How might $-\sum ln(p_i)$ be used to test H_0 vs H_1 using the large sample property?

HINT: By CLT, $\sqrt{n}(\frac{-\sum^{n} ln(P_i)}{n} - 1) \to_d N(0, 1)$

7 Exam 2005

7.1 Problem 1

TOPIC: Properties of Standard Normal and Function of Standard Normal r.v.s

Z_1, Z_2 are i.i.d. N(0, 1) r.v.s and define $Y_1 = Z_1 + Z_2, Y_2 = Z_2, W = exp(Y_1)$

Part a. Find $E[Y_1]$, Var$[Y_1]$, cov(Y_1, Y_2)

HINT: E$[Y_1]$ = 0, Var$[Y_1]$ = 1 + 1 = 2, cov(Y_1, Y_2) = Var(Z_2) = 1

Part b. Find the marginal density of Y_1

HINT: Uh well its the sum of two standard normals so $Y_1 \sim N(0, 2)$

Part c. Find the joint density function of Y_1, Y_2.

HINT: $Y_1 = Z_1 + Z_2, Y_2 = Z_2 \to Z_1 = Y_1 - Y_2, Z_2 - Y_2$. The Jacobian matrix for this is

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Plugging all of this into the original joint distribution function, we get $f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2\pi} exp(-\frac{(y_1-y_2)^2+y_2^2}{2})$

Part d. Find the density function of W.

HINT: Plugging into the formula for a univariate transformation, we should get that $f_w(w) = \frac{1}{w}\frac{1}{\sqrt{4\pi}} exp(-\frac{ln(w)^2}{4})$

Part e. Find E[W] and Var[W]

HINT: Use MGFs. $M_{Y_1}(t) = E[e^{Y_1 t}] = e^{t^2}$ so $E[W] = e$ and $Var[W] = E[e^{2Y_1}] - (E[e^{Y_1}])^2 = e^4 - e^2$

7.2 Problem 2

TOPIC: Inequalities

Let $\bar{X}_1$ and $\bar{X}_2$ be the sample means of two indpendent samples of size n from the same population with mean μ and variance σ^2.

Part a. Use the CLT to determine n s.t. the probability is about .01 that the two sample means will differ by more than σ

HINT:

$$P(\bar{X}_1 - \bar{X}_2 > \sigma) = .01 \to P(\frac{\bar{X}_1 - \bar{X}_2}{\sigma} > 1) = .01$$

Part b. Use Chebyshev's Inequality to determine n s.t. that probability that the means differ by more than σ is no greater than .01

HINT: Let Y = $\bar{X}_1 - \bar{X}_2$. Then $P(Y^2 \geq \sigma^2) \leq \frac{E[Y^2]}{\sigma^2} = \frac{2\sigma^2/n}{\sigma^2} = \frac{2}{n} = .01$

7.3 Problem 3

TOPIC: UMVUEs, MLEs with for Normal r.v.s with different variances

Let $X_1, \ldots, X_n$ be i.i.d. with $X_i \sim N(\theta, \frac{\sigma^2}{\lambda_i})$ where λ_i are known positive numbers and θ, σ^2 are unknown.

Part a. Find a CSS for (θ, σ^2)

HINT: Putting the pdf into exponential family form, we see that the CSS is $T = (\sum X_i, \sum X_i^2)$.

Part b. Find the UMVUE of θ

HINT: $E[\sum X_i] = n\theta$ so $W = frac\sum X_i n$ is the UMVUE of θ

Part c. Give a sufficient condition in terms of the λ_i for the consistency of the estimator.

HINT: We need the variance of our UMVUE to converge to 0. Doing the math, we see that $\frac{\sum \frac{1}{\lambda_i}}{n^2} \to 0$ is the necessary condition.

Part d. Find the MLE of σ^2

HINT: Differentiating the log-likelihood gives us $\hat{\sigma^2} = \frac{\sum \lambda_i (x_i - \theta)^2}{n}$.

7.4 Problem 4

TOPIC: Large sample Asymptotic tests with Bernoulli r.v.s

Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli(p) r.v.s where 0 ¡ p ¡ 1 is the unknown parameter. Find the following three asymptotic tests for testing $H_0 : p = 0.5$ vs $H_1 : p \neq 0.5$

Part a. Asymptotic likelihood ratio test

HINT: $\Lambda(p = .5) = -2nln(2) + ln(\bar{X})(\sum X_i) + ln(1 - \bar{X})(n - \sum X_i) \sim \chi_1^2$

Part b. Asymptotic Wald Test

HINT: $W(p_0) = \frac{(p_{MLE} - p_0)^2}{1/I_X(p_{MLE})} = \frac{n(\bar{X} - .5)^2}{\bar{X}(1 - \bar{X})} \sim \chi_1^2$

Part c. Asymptotic Score Test

HINT: $S(p_0) = \frac{U(p_0)^2}{I_X(p_0)} = \frac{(4\sum X_i - 2n)^2}{4n} \sim \chi_1^2$

7.5 Problem 5

TOPIC: UMP tests for Gamma r.v.s

$Y_1, \ldots, Y_n$ are i.i.d. observations from $f(y|\theta) = 2\theta\sqrt{\frac{\theta}{\pi}}\sqrt{y}exp(-\theta y), y > 0$

Part a. Derive a UMP test at level α of $H_0 : \theta \leq 1$ vs $H_1 : \theta > 1$

HINT: You can see that the monotone likelihood ratio here is increasing w.r.t. $\sum y_i$

Part b. Express the critical value k in tersm of a critical value from a known distribution

HINT: Since $Y_i \sim \Gamma(\frac{3}{2}, \frac{1}{\theta})$, $\sum Y_i \sim \Gamma(\frac{3n}{2}, \frac{1}{\theta})$

Part c. Describe how to calculate the power of the test in part (a) against the alternative $\theta = 2$

HINT: $P_{\theta=2}(\sum y_i < F^{-1}_{\Gamma(3n/2,1/2)}(\alpha))$

Part e. Find a function of T that is a pivot

HINT: $\theta \sum y_i \sim \Gamma(\frac{3n}{2}, 1)$

Part d. Use the pivot to derive a two sided confidence interval for θ

8 Exam 2006

8.1 Problem 1

TOPIC: Basic Properties of Probability Distributions

Consider $f(x) = cx^2, 0 < x < 1$ where c is some constant

Part a. Find the value of c

HINT: c = 3

Part b. Find the mean of the distribution

HINT: $E[X] = \frac{3}{4}$

Part c. Suppose X_1, X_2 are independent observations from the distribution. Let Y = $\min(X_1, X_2)$. Find $P(Y \leq .5)$

HINT: $P(Y \leq .5) = 1 - P(X_1 > y)P(X_2 > y) = 1 - (1 - .5^3)^2 = .234$

Part d. Find $E[Y]$

HINT: $f_y(y) = 6y^2 - 6y^5$ so $E[Y] = \frac{9}{14}$

Part e. Which is bigger, the mean of Y or the median of Y?

HINT: Just plug the mean into the CDF of Y

Part f. Let $Z = exp(X)$ where $X \sim f(x)$. Find the density of Z.

HINT: $f_z(z) = \frac{1}{z}3(ln(z))^2, -\infty < z < 0$

Part g. Is E[Z] smaller than, equal to, or bigger than 2.117?

HINT: Use Jensen's inequality. Since the exponential function is convex, $E[exp(X)] \geq exp(E[X]) = 2.117$

8.2 Problem 2

TOPIC: Tests on the Normal Distribution

$X_1, \ldots X_n \sim N(\mu_0, \sigma)$

Part a. Derive the LRT for testing $H_0 : \sigma = \sigma_0$ vs $H_1 : \sigma \neq \sigma_0$

Part b. Describe the asymptotic null distribution of the test statistic in part (a).

Part c. Suppose n = 20, $\sum X_i = 588$, $\sum X_i^2 = 17700, \mu_0 = 30, H_0 : \sigma = 3, H_1 : \sigma \neq 3$

Part d. Specify the exact distribution of the MLE of σ

HINT: $\frac{\sigma^2}{n} \sum \left(\frac{X_i - \mu_0}{\sigma}\right)^2 \sim \frac{\sigma^2}{n} \chi_n^2$

Part e. Use the distribution in (d) to obtain the test of $\sigma = \sigma_0$ vs $\sigma \neq \sigma_0$

8.3 Problem 3

Didn't do this one because **I didn't feel like it boiiii**.

8.4 Problem 4

Didn't do this one **either**.

8.5 Problem 5

TOPIC: Word Problem about Flowers!

There are 40 red flowers. Red flowers can have either an RR or RW genotype. Let θ be the proportion of RR red-flowered plants.

Our experiment is to have the 40 red flowers produce 2 offspring flowers apiece. For a RW parent, 50% of their offspring will be red and 50

Part a. If one red flower is collected at random, what is the probability that both offspring plants are red flowers.

Part b. What is the expected number of red flower offpsring in part (a)?

HINT: $E[X_i] = 0 + 1\frac{1}{2}(1-\theta) + 2(\frac{1}{4} + \frac{3}{4}\theta) = 1 + \theta$

Part c. Use MOM to obtain estimator of θ. Express as a function of Y_0, Y_1, Y_2 where Y_j is the number of plants that produced j red flowered plants.

Part d. Is your MOM unbiased?
HINT: haha yea.

9 Exam 2012

9.1 Problem 1

Suppose $W \sim N(0,1)$ and T_1, T_2 follow a distribution where $P(T = 1) = P(T = -1) = \frac{1}{2}$. Let X = WT_2 and $Y = WT_2$. W, T_1, T_2 are all independent.

Part a. Find the mgf of X.

Part b. Find Cov(X, Y)

Part c. Can (X, Y) have a joint bivariate normal distribution?

Part d. Let $Z = min(X, Y)$. Find the pdf of Z.

9.2 Problem 2

Z follows a Pareto distribution

$$f(x|a,\theta) = \begin{cases} \theta a^{\theta} x^{-(\theta+1)} & x > a \\ 0 & x \leq a \end{cases}$$

where $a, \theta > 0$.

Part a. Let $Y = ln\left(\frac{x}{a}\right)$. What is the distribution of Y?

Part b. From (a), provide a non-trivial lower bound for E[X] without calculating E[X] directly.

Part c. Find the MLE $\hat{a}$ and $\hat{\theta}$ using a random sample $X_1, \ldots, X_n$ i.i.d. Pareto(a, θ).

Part d. Prove that $\hat{a}$ and $\hat{\theta}$ are independent.

9.3 Problem 3

Suppose that $X_1, \ldots, X_n$ i.i.d Exp(θ) and $Y_1, \ldots, Y_n$ i.i.d Exp(2θ) and the Xs and independent from the Ys. $\theta > 0$ is unknown.

Part a. Find a complete sufficient statistic for θ.

Part b. Find the UMVUE of θ.

Part c. Find the UMVUE of θ^2.

Part d. SHow that the UMVUE of θ^2 is distributed independently of $\frac{\bar{Y}}{\bar{X}}$.

Part e. Derive the Fisher information contained in the entire dataset about θ.

9.4 Problem 4

$X_1, \ldots X_n \sim U(\theta - 1/2, \theta + 1/2)$, θ is unknown.

Part a. Show that $(X_{(1)}, X_{(n)})$ is jointly sufficient for θ.

Part b. Show that $\hat{\theta} = \frac{X_{(1)}+X_{(n)}}{2}$ is an unbiased estimator for θ.

Part c. Derive the CDF of $\hat{\theta}$ when $n = 2$.A

Part d. Find formulas for $P(\hat{\theta} > \theta + t)$ and $P(\hat{\theta} < \theta - t)$ with $0 < t < \frac{1}{2}$ and use them to derive an exact 90% two sided CI for θ when $n = 2$.

9.5 Problem 5

$X_1, \ldots X_n \sim$ Poisson(λ), $\lambda > 0$ and unknown. Let $T_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Part a. Given $\alpha \in (0, 1)$ and $\lambda_0 > 0$, derive a UMP test for $H_0 : \lambda = \lambda_0$ vs. $H_1 : \lambda > \lambda_0$.

Part b. Determine the asymptotic distribution of T_n as $n \to \infty$.

Part c. Given α, derive a large sample level α test for $H_0 : \lambda = \lambda_0$ vs. $H_1 : \lambda > \lambda_0$.

Part d. Find a variance stabilizing transformation $g(T_n)$ and determine its asymptotic distribution.

10 Exam 2013

I liked this exam a lot, relative to the other exams anyways :)

10.1 Problem 1

$X_1, \ldots X_n \sim U(-\theta, \theta)$, $\theta > 0$ is unknown.

Part a. Prove that $T = (X_{(1)}, X_{(n)})$ is a sufficient statistic for θ.

Part b. Is T minimal sufficient for θ?

Part c. Find the MLE of θ.

Part d. Prove that the MLE is a complete statistic for θ.

10.2 Problem 2

We have $X_1, \ldots, X_n \sim Gamma(\alpha_1, \beta_1)$ and $Y_1, \ldots, Y_n \sim Gamma(\alpha_2, \beta_2)$ where the X's and Y's are independent and all parameters are unknown.

Part a. Find the complete sufficient statistic (β_1, β_2).

Part b. Find the UMVUE for $\beta_2 - \beta_1$.

Part c. Find the UMVUE of $\frac{\beta_2}{\beta_1}$. Use the fact that if $Z \sim Gamma(\alpha, \beta)$, then $E[Z^r] = \beta^r \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)}$.

Part d. Show that the UMVUE of $\beta_2 - \beta_1$ is distributed independently of $\frac{Y_2}{Y_1}$.

Part e. Derive the Fisher information about (β_1, β_2) contained in the entire dataset.

10.3 Problem 3

Let X be a single distribution from a distribution with the following density

$$f(x|\theta) = \begin{cases} \frac{x}{\theta} & 0 \le x \le \theta \\ \frac{2-x}{2-\theta} & \theta \le x \le 2 \\ 0 & \text{o.w.} \end{cases}$$

for $\theta \in \Theta = (0, 2)$.

Part a. Find the critical function for the most powerful $\alpha = .05$ test of $H_0 : \theta = 1$ vs $H_1 : \theta = 0.1$. Specify the critical value.

Part b. What is the power of the test identified in (a) when $\theta = 0.1$?

Part c. Is the test in (a) unbiased for testing $H_0 : \theta = 1$ vs. $H_1 : \theta \neq 1$? If not, give an unbiased test.

Part d. Is there a UMP test at level $\alpha = .05$ of $H_0 : \theta = 1$ vs. $H_1 : \theta < 1$?

10.4 Problem 4

$X_1, \ldots X_n \sim$ Weibull$(2, \lambda)$ with pdf $f(x|\lambda) = \frac{2}{\lambda}\left(\frac{x}{\lambda}\right) exp\left(-\left(\frac{x}{\lambda}\right)^2\right), x > 0$.

Make use of the following facts:

$E[X_i] = \lambda\sqrt{\pi/4} = \mu$ $\qquad$ $Var[X_i] = \frac{4-\pi}{4}\lambda^2 = \frac{4-\pi}{4}\frac{4\mu^2}{\pi}$

Part a. What is the MLE of λ?

Part b. What is the information in a single observation of λ?

Part c. Construct the Wald statistic to test $H_0 : \lambda = \lambda_0$ vs. $H_1 : \lambda = \lambda_0$ at level $\alpha = .05$ and give the rejection region.

Part d. How could you use a standard t-test to test $H_0 : \lambda = \lambda_0$ vs. $H_1 : \lambda \neq \lambda_0$ at level $\alpha = .05$? Give the test statistic based on this sample of size n and the rejection region.

10.5 Problem 5

In this problem, you are asked to construct examples to illustrate statistical concepts?

Part a. Give an example where two random variables X and Y are uncorrelated but not independent.

Part b. Give an example of two sequences random variables $\{X_i\}$ and $\{Y_i\}$ which converge in distribution to X and Y but $\{X_i + Y_i\}$ does not converge to $X + Y$.

Part c. Consider an infinite sequence of i.i.d. normal random variables with known μ and variance $= 1$. Construct a sequence of estimators T_n such that T_n is an unbiased estimator of μ but the sequence T_n is not consistent for μ.

Part d. Give an example of a family of parametric distributions which has a minimal sufficient statistic that is not complete sufficient.

11 Exam 2014

11.1 Problem 1

$X_1, \dots X_n \sim$ Bernoulli(p). Consider $\beta = \log\left(\frac{p}{1-p}\right)$.

Part a. Prove that $P(\bar{X} = 0) \to 0$ and $P(\bar{X} = 1) \to 0$ as $n \to 0$.

Part b. Show that $\hat{\beta} = \log\left(\frac{\bar{X}}{1-\bar{X}}\right)$ is a consistent estimator of β.

Part c. Prove that $\sqrt{n}(\hat{\beta} - \beta) \to N(0, \frac{1}{p(1-p)})$ in distribution.

Part d. Show that $\hat{\beta}$ is asymptotically efficient by comparing its asymptotic variance to the CRLB.

11.2 Problem 2

$X_1, \dots, X_n$ be i.i.d. $f(x|\theta) = exp(-(x-\theta)), x > \theta$ and θ is unknown.

Part a. Show that the distribution family has MLR in some statistic T.

Part b. Show that the minimum order statistic $X_{(1)}$ has pdf $f_{X_{(1)}}(x|\theta) = ne^{-n(x-\theta)}$

Part c. For a given $0 < \alpha < 1$, derive a UMP test at level α for $H_0 : \theta = \theta_0$ vs $H_1 : \theta > \theta_0$

Part d. What is the MLE of θ?

Part e. What is the asymptotic distribution of the MLE found in (d)?

11.3 Problem 3

Let A and B be two different events related to a random experiment. Suppose that n independent trials of the experiment are carried out and the frequenceis of the occurrences of the events are given as:

	A	A^C
B	n_{11}	n_{12}
B^C	n_{21}	n_{22}

where $n = n_{11} + n_{12} + n_{21} + n_{22}$. Consider testing the following hypotheses $H_0 : P(A) = P(B)$ vs $H_1 : P(A) \neq P(B)$.

Part a. Let $p_{11} = P(A \cap B)$, $p_{12} = P(A^C \cap B)$, and $p_{21} = P(A \cap B^C)$. Derive the unconstrainted MLEs of p_{11}, p_{12}, p_{21}.

Part b. Assume that $P(A) = P(B)$. Derive the constrained MLEs of p_{11}, p_{12}, p_{21}.

Part c. Derive the LRT $\Lambda(x)$ for testing $H_0 : P(A) = P(B)$ vs $H_1 : P(A) \neq P(B)$.

Part d. Given a reasonable α, derive a large sample level α test based on the asymptotic distribution of $-2ln(\Lambda)$.

Part e. Suppose we observe that $n_{11} = 13, n_{12} = 36, n_{21} = 39, n_{22} = 12$. Compute the test statistic and the associated p-value.

11.4 Problem 4

$X_1, \ldots, X_n$ be i.i.d. $f(x|\alpha, \theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{-\alpha-1} exp\left(-\frac{1}{\theta x}\right)$. This is the inverse Gamma distribution.

Use the following facts:

$E[X_i] = \frac{1}{\theta(\alpha-1)}, \alpha > 1$	$Var[X_i] = \frac{1}{\theta^2(\alpha-1)^2(\alpha-2)}$
$E[1/X_i] = \theta\alpha$	$Var[1/X_i] = \theta^2\alpha$
	$E[exp(-\sum ln(x_i))] = (\alpha\theta)^n$

Part a. Find the minimum sufficient statistic T for (α, θ).

Part b. Find the MLE for θ, assuming that α is known. Is the MLE a consistent estimator for θ?

Part c. Find the method of moments estimator for θ, assuming that α is known. Is the MOM estimator consistent for θ?

Part d. Find the UMVUE for θ, assuming α is known.

Part e. Find the CRLB for the variance of unbiased estimators of θ, assuming α is known. Does the variance of the UMVUE achieve the bound?

11.5 Problem 5

Let $X = (X_1, \ldots, X_n)$ be a sample of random variables with joint cumulative distribution function F_X which depends on unknown parameter θ. Let $T(X)$ be an unbiased estimator for θ and $S(X)$ be a sufficient statisitc for θ. Define $W = E[T|S]$.

Part a. Show that W is an estimator.

Part b. Show that $E[W] = \theta$ so W is an unbiased estimator for θ.

Part c. Show that $Var(W) \leq Var(T)$ so W is more efficient than T.

Part d. Recall Markov's Inequality. If X is a nonnegative random variable $P(X > 0) = 1$ with finite expected value $E[X] = \mu$.

Part e. Let $X_1, \ldots, X_n$ be i.i.d. sample from a distribution with cdf $F(x) = P(X_i \leq x)$. Prove that $\hat{F}(x) = \frac{1}{n} \sum^n I_{x_i < x}$ is a consistent estimator for $F(x)$. What is the asymptotic distribution of $\sqrt{n}(\hat{F}(x) - F(x))$?

12 Exam 2015

12.1 Problem 1

$X_1, \ldots X_n \sim Exp(\beta)$.

Part a. Use Markov's Inequality to give an upper bound of $P(X_1 \geq 10)$

Part b. Determine the pdf of X_1^2.

Part c. State the exact distribution of $\bar{X}$.

Part d. Determine the asymptotic distribution of $\bar{X}$.

12.2 Problem 2

Suppose we have $X_1, \ldots, X_n \sim F$, $Y_1, \ldots, Y_n \sim G$ with F and G the unknown. Xs and Ys are independent and we are interested in the following parameter $p = P(X + Y > 0)$. We will consider various estimates of p.

Part a. Consider $\hat{p} = \frac{1}{n} \sum_{i=1}^n I_{x_i + y_i > 0}$. Find the bias and MSE of $\hat{p}$.

Part b. Consider the estimator $\hat{p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n I_{x_i + y_i > 0}$. Find the bias and MSE of $\hat{p}$.

Part c. Suppose $F = N(\mu_1, \sigma_1^2)$ and $G = N(\mu_2, \sigma_2^2)$ where all parameters are unknown. Find the MLE p^* of p.

Part d. Find the asymptotic distribution of the MLE p^* as $n \to \infty$.

12.3 Problem 3

$X_1, \ldots, X_n \sim Exp(\theta)$ which has $f(x|\theta) = \frac{1}{\theta} exp\left(-\frac{x}{\theta}\right), x > 0, \theta > 0$

Part a. Identify a pivotal quantity U for θ based on a sufficient statistic for θ. Use this pivot to find an equal-tail $1 - \alpha$ CI for θ.

Part b. Find the MLE for θ based on the sample and provide the asymptotic distribution as $n \to \infty$.

Part c. Use your results from (b) to give an asymptotic $1-\alpha$ Wald Confidence interval for θ. Transform this interval to obtain an asymptotic $1 - \alpha$ CI for θ^2.

Part d. Find the asymptotic distribution of $\hat{\theta}^2_{MLE}$ as $n \to \infty$.

Part e. Use your results from (d) to give an asymptotic $1 - \alpha$ Wald CI for θ^2. Compare the length of the interval found in (c).

12.4 Problem 4

Let X be one observation from the discrete distribution with pmf $f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1 - \theta)^{1-|x|}, x = -1, 0, 1.$ $0 < \theta < 1$.

Consider the statistics $T_1 = X$ and $T_2 = |X|$.

Part a. Does $f(x|\theta)$ belong to the exponential family?

Part b. Is T_1 sufficient for θ? What about T_2? (Man, what a great movie)

Part c. Is T_1 minimum sufficient? What about T_2?

Part d. Using the definition, determine whether T_1 or T_2 is complete.

Part e. Is the statistic $T_3 = X^2$ a complete sufficient statistic? Explain.

12.5 Problem 5

Let f and g be two known pdfs on $\mathbf{R}$. Let X be a single observation from this pdf.

Furthermore, define a mixture distribution $\theta f(x) + (1 - \theta) g(x)$ where $0 < \theta < 1$ is unknown. α is a given constant defining the size of a test.

Part a. Consider $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$. Find a UMP test of size α.

Part b. Suppose $\psi(x)$ is the critical function. Show that the power function is a linear funciton of θ.

Part c. Consider $H_0 : \theta = \theta_0$ and $H_1 : \theta \neq \theta_0$ with $0 < \theta_0 < 1$. Show that no UMP test exists.

Part d. Given $0 < \theta_1 < \theta_2 < 1$, consider $H_0 : \theta \leq \theta_1$ or $\theta \geq \theta_2$ vs $H_1 : \theta_1 < \theta < \theta_2$. Use the conclusion from (b) to show that $\psi(x) = \alpha$ is a UMP test of size α.

13 Exam 2016

This exam was long! Most people also considered this exam fairly difficult.

13.1 Problem 1

Suppose that $X_1 \sim N(0, 1)$. In each of the following scenarios, $X_2|X_1$ is given and a question is asked.

Part a. Suppose that $X_2|X_1 = x_1 \sim N(x_1, \sigma^2)$. Show that $\text{Cov}(X_2, X_1) = 1$.

Part b. Suppose that $X_2|X_1 = x_1 \sim N(x_1^2, \sigma^2)$, $\forall x_1 \in \text{R}$.
Show that $\text{Cov}(X_2, X_1) = 0$.

Part c. Suppose that $X_2|X_1 = x_1 \sim N(0, x_1^2)$, $\forall x_1 \in \text{R}$. Find the mgf of X_2.

Part d. Suppose that $X_2|X_1 = x_1 \sim N(x_1, \sigma^2)$, $\forall x_1 \in \text{R}$. Find the marginal distribution of X_2.

Part e. Suppose that $X_2|X_1 = x_1 \sim N(x_1, x_1^2)$, $\forall x_1 \in \text{R}$. Show that $U = X_1$ and $V = \frac{X_2}{X_1}$ are independent and find the marginal distribution of V.

13.2 Problem 2

Suppose that $X_1, \ldots, X_n$ i.i.d. with pmf $p(x|\theta, \alpha) = \frac{\theta^x (1+\alpha x)^{x-1} e^{-\theta(1+\alpha x)}}{x!}$ where $\theta > 0$ and $0 < \alpha < \frac{1}{\theta}$. The population mean and variance for this distribution are $\mu = \theta(1 - \alpha\theta)^{-1}$ and $\sigma^2 = \theta(1 - \alpha\theta)^{-3}$.

Part a. If α is known, find a minimal sufficient statistic for the parameter θ.

Part b. If α is known, find the MLE of θ.

Part c. If α is known, find the asymptotic distribution of the MLE of θ.

Part d. If α is known, find the method of moments estimator of θ.

Part e. Assuming both parameters are unknown, find the method of moments estimator for α and θ.

13.3 Problem 3

Suppose that $X_1, \ldots, X_n$ are i.i.d. $f(x|\theta, \lambda) = \frac{1}{\lambda} exp\left(-\frac{(x-\theta)}{\lambda}\right) I_{x>\theta}$.

Part a. Find the MLE of (θ, λ).

Part b. Find the MLE of $(\theta, log(\lambda)$.

Part c. Find the UMVUE of $log(\lambda)$, assuming that θ is known.

Part d. Consider θ known. What is the CRLB of $\log(\theta)$? Does the UMVUE of $\log(\theta)$ attain the CRLB?

HINT: This problem was deceptively hard. We had to use Corollary 7.3.15 from Casella Berger.

Part e. Is the θ component of the MLE from part b. a consistent estimator of θ?

Part f. Is the $\log(\theta)$ component of the MLE from part b. a consistent estimator of $\log(\lambda)$?

13.4 Problem 4

Consider $U_1, \ldots, U_n$ and $W_1, \ldots, W_n$ are i.i.d. $N(0, \sigma^2)$ and suppose that we can only observe the transformed variables $X_i = \sqrt{U_i^2 + W_i^2}, i = 1, \ldots, n$.

Part a. Show that X_i follows a Rayleigh distribution with density $f(x|\sigma^2) = \frac{x}{\sigma^2} exp\left(-\frac{x^2}{2\sigma^2}\right)$ where $x \geq 0$.

Part b. Find the MLE $\hat{\sigma^2}$ of σ^2 based on $X_1, \ldots, X_n$.

Part c. Find the asymptotic distribution of $\hat{\sigma^2}$ in (b).

Part d. Consider testing the hypothesis that $H_0 : \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$. Obtain a large sample α test based on the asymptotic distribution of $\hat{\sigma^2}$.

Part e. Find the power function of the test in (d).

Part f. What is the limit of the power function as $n \to \infty$.

13.5 Problem 5

Consider 3 independent Bernoulli samples of size n with $X_{i1}, \ldots, X_{in}$ i.i.d. Bernoulli(p_i).

p_1, p_2, p_3 are unknown mean parameters.

Part a. Find the UMP level α test for $H_0 : p_1 = .5$ vs $H_1 : p_1 > .5$ based on the first sample $X_{11}, \ldots, X_{1n}$.

Part b. Consider

$$\bar{X}_j = \begin{pmatrix} \frac{1}{n}\sum_{j=1}^{n} X_{1j} \\ \frac{1}{n}\sum_{j=1}^{n} X_{2j} \\ \frac{1}{n}\sum_{j=1}^{n} X_{3j} \end{pmatrix}$$

with $j = 1, \ldots, n$.

Determine the asymptotic distribution of $\bar{X}_j$.

Part c. Find a 95% large sample joint confidence set for (p_1, p_2, p_3).

Part d. Find a large level α test for testing $H_0 : \frac{p_1+p_2}{2} = p_3$ vs $H_1 : \frac{p_1+p_2}{2} \neq p_3$.

Part e. Find a large sample level α test for $H_0 : p_1 = p_2 = p_3$ vs. $H_1 : p_1, p_2, p_3$ not all equal.

14 Exam 2017

This exam was notoriously hard! Keep that in mind as you are working through it.

14.1 Problem 1

Let X and Y be independent r.v.s following an exponential distribution with means a and b respectively. Define $Z = min(X, Y)$ and

$$W = \begin{cases} 1 & \text{if } X \leq Y \\ 0 & \text{if } X > Y \end{cases}$$

Part a. Find $P(W = 1)$

HINT: Start with joint pdf of X and Y and then use that to calculate $P(X \leq Y)$. Should get $\frac{a}{a+b}$ as a final answer.

Part b Find the pdf of Z

HINT: Start with calculating the cdf of Z (aka $1 - P(Z > z)$) then take derivative.

Part c. Find the joint distribution of W and Z (i.e. $P(Z \leq z, W \leq w)$)

HINT: Calculate $P(Z \leq z, W = w)$ using law of total probability by breaking it up into two parts, one conditioned on $W = 0$ and another conditioned on $W = 1$.

Part d. Are W and Z independent?

HINT: Yes, should be able to factor $P(Z \leq z, W = 0)$ and $P(Z \leq z, W = 1)$ apart into separate pieces for W and Z.

Part e. Find $E[X|W = 1]$

HINT: Calculate $P(X|W = 1)$ first, then calculate the integral for $E[X|W = 1]$ (this integral should be able to be quickly solved using gamma kernel trick).

14.2 Problem 2

Given $\theta > 0$, consider the following three scenarious where $X_1, \ldots, X_n$ i.i.d. with different pdfs:

i. $U(0, \theta)$
ii. $U(-\theta, \theta)$
iii. $U(\theta, 2\theta)$

Part a. Find the method of moments estimator and MLE of (i), (ii), and (iii).

Part b. For case (i), determine which estimator should be preferred in terms of MSE.

Part c. For case (ii), find the best unbiased estimator of θ.

Part d. For case (iii), determine which estimator can be improved by using sufficiency.

14.3 Problem 3

Let $X_1, \ldots, X_n$ i.i.d $f(x|\theta) = \frac{\theta}{x^{\theta+1}} I_{x>1}$ where $\theta > 0$ is unknown.

Part a. Find the MLE of θ

Part b. Find the UMVUE of θ. Which of the two estimators of θ, the MLE or the UMVUE, is the better estimator in terms of MSE? Use fact that if $Y \sim Gamma(\alpha, \beta)$, then $E[Y^{-1}] = \frac{1}{(\alpha-1)\beta}$ and $V[Y^{-1}] = \frac{1}{(\alpha-1)^2(\alpha-2)\beta^2}$

Part c. What is the CRLB for unbiased estimators of θ? Does the UMVUE of θ attain the CRLB?

Part d. Show that as $n \to \infty$, $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} ln(x_i) - \frac{1}{\theta}\right) \to^d N(0, \frac{1}{\theta^2}$

Part e. Find the asymptotic distribution of the geometric mean $\left(\prod_{i=1}^{n} X_i\right)^{1/n}$.

14.4 Problem 4

Let $X_1, \ldots, X_n$ i.i.d. $U(0, \theta), \theta > 0$. Consider the maximum statistic $X_{(n)}$.

Part a. Show that the statistic $X_{(n)}$ is consistent.

Part b. Is $X_{(n)}$ unbiased? If not, find an unbiased estimator based on $X_{(n)}$.

Part c. Find the asymptotic distribution of $X_{(n)}$. Use this result to compute a confidence interval with appropriate confidence level $100(1-\alpha)$

Part d. Show that the family of distributions $U(0, \theta) : \theta > 0$ has monotone likelihood ratio and construct a UMP test (if it exists) of size α for testing $H_0 : \theta \leq \theta_0$ vs. $H_A : \theta > \theta_0$

14.5 Problem 5

This problem was pretty ridiculous so I ended up not doing it.

15 Exam 2018

This exam was pretty universally regarded as fair, so I would recommend doing this exam a day or two before your actual comprehensive exams as a warm up.