Statistics 24400 - Autumn 2016

Final Examination Solution

December 2 and 5, 2016

Name (print): ______________________________

On my honor, I will not discuss this exam with **ANY PERSON** before 15:30 December 5, 2016.

Signature ______________________________

1. Please **print** your name in the space provided. If you are taking this exam on the 2nd or the morning of the 5th, you must sign the "temporary nondisclosure" line and conduct yourself accordingly in order to get credit for this final exam.

2. Do not sit directly next to another student.

3. Do not turn the page until told to do so.

4. This is a closed book examination. You are allowed a single page of notes, written on both sides. Please write your name on your notes and turn it in with the exam. You are permitted to have a calculator. Devices capable of communication (laptops, tablets, phones) must be powered down. Tables of the cumulative Normal and χ^2 distributions are at the end of the exam.

5. Please provide the answers in the space and blank pages provided. If you do not have enough space, please use the back of a nearby page, clearly indicating the identity of the continued problem.

6. Be sure to show your calculations. In order to receive full credit for a problem, you must show your work and explain your reasoning. Good work can receive substantial partial credit even if the final answer is incorrect.

7. Read through the exam before answering any questions. Our scale of credit for questions may not correlate with the level of difficulty you experience—use your time wisely!

Question	*Points*	*Score*
Question 1	20	
Question 2	30	
Question 3	30	
Question 4	20	
TOTAL	100	

1. **True or False?** (20 pts) There is a "guessing penalty" penalty of four questions (8 pts). You'll begin to accrue points with your **fifth** correct answer.

   (a) (2 pts) T F
   If random variables X and Y are uncorrelated (i.e. $\mathrm{Cor}(X,Y)=0$), then X and Y must be independent. **F**

   (b) (2 pts) T F
   If $P(A) < P(B)$ and $P(C) > 0$, then $P(A|C) \le P(B|C)$. **F**

   (c) (2 pts) T F
   The calculation of p-value does not depend on the alternative hypothesis once we know the null hypothesis. **T**

   (d) (2 pts) T F
   The maximum likelihood estimator (MLE) is always unbiased. **F**

   (e) (2 pts) T F
   For any hypothesis testing procedure, it is always possible to increase the power $\pi = 1-\beta$ while keeping the type 1 error α the same. **F**

   (f) (2 pts) T F
   If X is a continuous random variable with cdf $F(X)$, and Y is a random variable such that $Y = F(X)$, then Y is distributed uniformly on $[0,1]$. **T**

   (g) (2 pts) T F
   $\Gamma(\frac{3}{2}) < \Gamma(\frac{1}{2})$ **T**

   (h) (2 pts) T F
   Fisher's method of combination on a set of tests of the same hypothesis with p-values $p_1, p_2, \ldots p_k$ means that the p-value of the combined tests is given by $P = p_1 p_2 \ldots p_k$. **F**

   (i) (2 pts) T F
   P defined as above. Then $-2\log P \sim \chi^2_{2k}$. **T**

   (j) (2 pts) T F
   Suppose that $X_i \sim \mathcal{N}(\mu, \sigma^2)$, with μ and σ unknown. Then the distribution of $X_1 - \mu$ does not depend on any unknown parameter. **F**

   (k) (2 pts) T F
   As above, suppose that $X_i \sim \mathcal{N}(\mu, \sigma^2)$ with μ and σ unknown, Then the distribution of $\frac{X_1 - X_2}{X_3 - X_4}$ does not depend on any unknown parameter. **T**

   (l) (2 pts) T F
   For a Poisson process, conditional on the number of events $N(0,1] = n$, the number of events
   $$N(0,1/3] \sim \mathrm{Bin}(n, 1/3)$$
   T

   (m) (2 pts) T F
   Fisher's Approximation can always be used to find the asymptotic variance of a Maximimum Likelihood Estimator. **F**

(n) (2 pts) T F
The proof of the Neyman-Pearson Lemma involves a Taylor expansion. **F**

2. **The Asymptotic Distribution of an MLE** (30 pts)
The density of the Weibull distribution is

$$f(x) = \frac{\beta}{\alpha^\beta} x^{\beta-1} \exp\{-(x/\alpha)^\beta\} \text{ for } x > 0$$

Assume that β is known, and we want to estimate α.

(a) (10 pts) Show that if W follows the Weibull distribution with parameters α and β, then $X = W^\beta$ follows an exponential distribution with parameter λ. Write an equation that gives λ in terms of α.

(b) (10 pts) Find the MLE for α in terms of n iid transformed observations (that is, in terms of the X_i's, where $X_i = W_i^\beta$ and the W_i's are the original observations). If you can't do the first part, you'll get partial credit by finding the MLE for λ.

(c) State the distribution of the MLE for large n, and calculate its mean and variance.

3. **Asymptotic Distribution of MLE**

(a) W follows Weibull distribution with parameters α and β and we consider $X = W^\beta$. So, the density of X is

$$\begin{aligned} f_X(x) &= f_W(x^{1/\beta}) \left| \frac{\partial}{\partial x} x^{1/\beta} \right| \\ &= \frac{\beta}{\alpha^\beta} x^{1-1/\beta} \exp\{-x/\alpha^\beta\} \mathbf{1}\{x > 0\} \times \frac{1}{\beta} x^{1/\beta - 1} \\ &= \frac{1}{\alpha^\beta} \exp\{-x/\alpha^\beta\} \mathbf{1}\{x > 0\} \end{aligned}$$

Clearly, X follows an exponential distribution with mean parameter α^β. In other words, $X \sim$ Exponential (λ) where $\lambda = \alpha^{-\beta}$.

(c) Now, we have to find the mle for α based on n iid transformed observations X_i, which can be thought to be coming from exponential distribution with parameter $\lambda = \alpha^{-\beta}$. We have already seen in the lecture that $\hat{\lambda} = 1/\bar{X}$. So, substituting original values, we can get that

$$\boxed{\hat{\alpha}_{\text{mle}} = \left(\frac{1}{n} \sum_{i=1}^{n} w_i^\beta \right)^{1/\beta}}$$

(d) Finally, we have to find the asymptotic distribution of these mle's for large n. So, we consider the statistic

$$\hat{\alpha} = \left(\frac{1}{n} \sum_{i=1}^{n} X_i \right)^{1/\beta}$$

Now, recall that if $\hat{\theta}_n$ is the mle of θ and the true value is θ_0, then $\sqrt{nI(\hat{\theta}_0)}(\hat{\theta}_n - \theta_0)$ converges to $N(0,1)$ random variable. Here, if $X_1, \ldots, X_n$ are coming from Exponential (λ) distribution, then, the information is

$$\begin{aligned} I(\alpha) &= \mathbb{E}\left(-\frac{\partial}{\partial \alpha^2} \log L(\alpha; X_1) \right) \\ &= \mathbb{E}\left(-\frac{\partial}{\partial \alpha^2} \left[-\beta \log \alpha - \frac{X_1}{\alpha^\beta} \right] \right) \\ &= \mathbb{E}\left(-\frac{\partial}{\partial \alpha} \left[-\frac{\beta}{\alpha} + X_1 \beta \alpha^{-\beta-1} \right] \right) \\ &= \mathbb{E}\left(-\frac{\beta}{\alpha^2} + X_1 \beta(\beta+1) \alpha^{-\beta-2} \right) \\ &= -\frac{\beta}{\alpha^2} + \frac{\beta(\beta+1)}{\alpha^2} = \frac{\beta^2}{\alpha^2} \end{aligned}$$

Hence, the asymptotic distribution should be of the form

$$\frac{\sqrt{n}(\hat{\alpha} - \alpha)}{(\alpha/\beta)} \Rightarrow N(0, 1)$$

4. **Sufficient Statistics and the Normal Distribution** [30 pts]

Let $X_1, X_2, \cdots, X_n$ be n i.i.d. samples from a $N(\mu, \sigma^2)$ distribution .

(a) Assume in this subpart and (b) that $\sigma = \sqrt{\mu}$ and $\mu > 0$. Give a sufficient statistic (apart from the data itself). [5 pts]

(b) Will the MLE be a function of the sufficient statistic in (a)? Justify your answer. [5 pts]

(c) Assume now that $\mu = 0$ and σ is unknown. Derive the MLE for σ (not σ^2). Is this estimator unbiased? [5 + 5 = 10 pts]

(d) Now assume that σ^2 is known but μ is not known. A Likelihood Ratio Test for $H_0 : \mu = \mu_0$ and $H_1 : \mu \neq \mu_0$ is a test that rejects H_0 when $\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} > c$. As a statistical investigator, you want a Type I error of 0.05 and Type II error of 0.25 when $\mu = \mu_0 + \sigma$. Find the values of n and c that will achieve this. [10 pts]

Solution

1.Substituting $\sigma^2 = \mu$, we get,

$$\frac{f(x|\mu,\mu)}{f(y|\mu,\mu)} = \exp\left[\sum_i x_i - \sum_i y_i\right] \exp\left[-\frac{1}{2\mu}\left(\sum_i x_i^2 - \sum_i y_i^2\right)\right]$$

This is constant as a function of μ iff $\sum_i x_i^2 = \sum_i y_i^2$. Thus $\sum_i X_i^2$ is the minimal sufficient statistic.

2.The likelihood can be written as

$$L(\mu,\sigma|X_1, X_2, \cdots, X_n) = h(X_1, X_2, \cdots, X_n) g(\mu, \sigma|T(X_1, X_2, \cdots, X_n))$$

where T is the sufficient statistic.
Then

$$\log L(\mu,\sigma|X_1, X_2, \cdots, X_n) = \log h(X_1, X_2, \cdots, X_n) + \log g(\mu, \sigma|T(X_1, X_2, \cdots, X_n))$$

So differentiating with respect to μ and σ, the function L is equivalent to differentiating g and that is a function of the sufficient statistic only. So, the solution (which is the MLE) will be a function of the sufficient statistic.

3.

$$L(\sigma|X_1, X_2, \cdots, X_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-X_i^2/2\sigma^2\right)$$

$$\log L(\sigma|X_1, X_2, \cdots, X_n) = C - n\log\sigma - \frac{\sum_i X_i^2}{2\sigma^2}$$

$$\frac{\delta \log L}{\delta\sigma} = -\frac{n}{\sigma} + \frac{\sum_i X_i^2}{\sigma^3} = 0$$

By solving this we get

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$$

$$\frac{\delta^2 \log L}{\delta\sigma^2} = -3\frac{\sum_i X_i^2}{\sigma^4} + \frac{n}{\sigma^2}$$

This is negative at $\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$.
So, the MLE of σ is $\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$.
The estimator is not unbiased for σ because by Cauchy Schwartz inequality,

$$E(Y^2) \geq E(|Y|)^2$$

$$\sigma^2 = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)$$

$$\sigma = \sqrt{E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right)} \geq E\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}\right)$$

So the estimator $\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$ is not unbiased for σ.

4.The power function

$$\beta(\theta) = Pr_\theta\left(\frac{|\bar{X} - \theta_0|}{\sigma/\sqrt{n}} > c\right)$$

By location and scale changes, it can be written as

$$\beta(\theta) = 1 + \Phi\left(-c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right) - \Phi\left(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)$$

where Φ is the normal CDF.

$$\beta(\theta_0) = 0.05 = 1 + \Phi(-c) - \Phi(c)$$

This when solved gives $c = 1.96$. The power (1- Type II error) then corresponds to

$$0.75 = \beta(\theta_0 + \sigma) = 1 + \Phi\left(-c - \sqrt{n}\right) - \Phi\left(c - \sqrt{n}\right) \approx 1 - \Phi(1.96 - \sqrt{n})$$

as the second term would be small. From here we get

$$1.96 - \sqrt{n} = -0.675$$

This implies

$$n \approx 6.9 \approx 7$$

5. **Grades From Scary Professors** (20 pts)
Three professors at the University of R'lyeh, Cthulhu, Azathoth, and Yog-Sothoth, give a statistics exam. The breakdown of grades for each professor is as follows:

Grade	Prof. Cthulhu	Prof. Azathoth	Prof. Yog-Sothoth	Total
A	15	30	20	65
B	35	40	45	120
C	30	25	35	90
D or F	10	5	10	25
Total	90	100	110	300

We are interested in determining whether there are systematic difference of grading policy among the three professors.

(a) (5 points) Formulate a suitable null and alternative hypothesis. Is this a test of independence, a test of homogeneity, or something else? Explain briefly.

(b) (9 points) Formulate and calculate a chi-squared test. Do you reject the null hypothesis in (a) at the 5% level of significance?

(c) (3 points) Of all the entries in the above table, which one contains the strongest evidence that there might, indeed, be a difference among the professors grading policies?

(d) (3 points) Briefly discuss and interpret your answers to (b) and (c).

Solution

(a) If we use symbols like $P_C(a)$ as an abbreviation for "the probability that Cthulhu gives an A", then the natural null hypothesis is

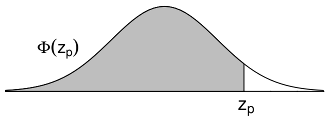
$$H_0 : P_C(a) = P_A(A) = P_Y(A), P_D(B) = P_A(B) = P_Y(B),$$
$$P_D(C) = P_A(C) = P_Y(C), P_D(D/F) = P_A(D/F) = P_Y(D/F)\,.$$

and H_1 is that at least one of the equalities in H_0 is false. This is a test of homogeneity: the column totals (number of students in each class) is fixed, but the hypothesis amounts to the claim that the conditional probability of getting a certain grade, given which professor, is the same for all three professors.

(b) The chi-squared statistic is 8.44, with $3\times 2 = 6$ degrees of freedom. We reject H_0 at 5 level of significance if the 0.05 upper quantile of χ^2 distribution with 6 degrees of freedom is less that 8.44. But from the "given Table", we find that the probability that a χ^2 random variable with 6 degrees of freedom is larger than 8.44 lies between 0.25 and 0.1. Thus the result is not statistically significant at 5 level of significance.

(c) The largest single entry in the table of $\frac{(f_o - f_e)^2}{f_e}$ is 3.21, corresponding to the number of A's given by Professor Azathoth.

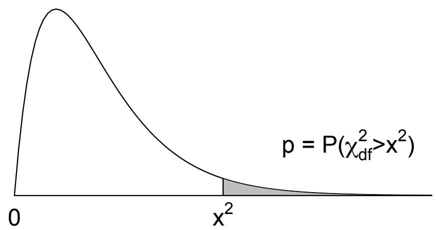
(d) The table suggests some differences among the three professors, in particular, Azathoth seems to give more A's than the other two professors. However, the fact that the chi-squared test led did not reject the null hypothesis implies that the differences could be random – there is no systematic evidence of a discrepancy in grading policy.



Cumulative Normal Distribution

z	**.00**	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Also for $z_p = 4.0$, 5.0, and 6.0, the values of p are 0.99997, 0.9999997, and 0.999999999.



χ^2 Distribution

df	Tail probability p											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.52	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2