Ph.D. Qualifying Examination in Statistics

Department of Mathematics
University of Louisville
August 12, 2016, 9:00am–12:30pm

Do 3 problems from each section.

SECTION 1

Problem 1. Suppose Y is a random variable with probability density function (pdf)

$$f(y|\alpha) = \begin{cases} \dfrac{\alpha - 1}{y^{\alpha}} & \text{for } y > 1 \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 1$.

(a) A family of probability density functions is called an exponential family if it can be expressed as $f(x|\theta) = h(x)C(\theta)\exp\{W(\theta)t(x)\}$. Is $\{f_Y(y|\alpha)\}$ an exponential family? If yes, define θ and find $h(y)$, $C(\alpha)$, $W(\alpha)$, and $t(y)$? If not, justify your answer.

(b) Find $E[\ln Y]$.

Problem 2. Suppose $X_1, \ldots, X_n$ is a random sample from a normal distribution with mean 0 and variance θ^2 having probability density function

$$f(x|\theta) = \frac{1}{\sqrt{2\pi\theta^2}} \exp\left\{-\frac{1}{2\theta^2}x^2\right\}, \theta > 0.$$

(a) Find the maximum likelihood estimator of θ^2.

(b) Is the estimator in part (a) sufficient for θ^2? Justify your answer.

(c) What is the maximum likelihood estimator of $P(X_1 > \theta^2)$? Write your answer in terms of the standard normal cdf $\Phi(t) = \displaystyle\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}u^2\right\} du$?

Problem 3. Let X be a Poisson(λ) random variable having probability mass function

$$P(X = x) = \begin{cases} \dfrac{1}{x!}\lambda^x e^{-\lambda} & \text{if } x = 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Suppose that $X_1, \ldots, X_n$ are independent Poisson(λ) random variables.

(a) Show that $\bar{X} = \sum_{i=1}^{n} X_i/n$ is an unbiased estimator of λ.

(b) What is the variance of $\bar{X}$?

(c) Show that $\bar{X}$ satisfies the Cramér-Rao Lower Bound for an unbiased estimator of λ.

Problem 4. Suppose that $X_1, \ldots, X_n$ are independently and identically distributed according to the uniform$(0, \theta)$ distribution. Let $M_n = \max(X_1, \ldots, X_n)$. Let

$$\delta_c(X_1, \ldots, X_n) = \begin{cases} 1 & \text{if } M_n \geq c \\ 0 & \text{otherwise} \end{cases}$$

be the test function for which $\delta_c = 1$ indicates that the null hypothesis should be rejected and $\delta_c = 0$ indicates that it should not be rejected.

(a) For $0 < c < \theta$, compute the power function $\beta(\theta, \delta_c) = P_\theta\left(\delta_c = 1\right)$.

(b) In testing $H_0 : \theta \leq \frac{1}{2}$ versus $H_A : \theta > \frac{1}{2}$, what choice of c would make the test based on the test function δ_c have size 0.05? Your choice of c should depend on n.

(c) How large should n be so that the test specified in (b) has power 0.98 for $\theta = \frac{3}{4}$?

SECTION 2

Problem 5. Let $\boldsymbol{y}$ be an n-dimensional column vector, $\boldsymbol{X}$ be an $n \times k$ matrix, $\boldsymbol{\beta}$ be a k-dimensional column vector, and $\mathbf{0}_k$ is a k-dimensional vector of zeros.

(a) Let $Q(\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2$. Compute $\dfrac{\partial Q}{\partial \boldsymbol{\beta}}$.

(b) Derive the solution to the score equation $\dfrac{\partial Q}{\partial \boldsymbol{\beta}} = \mathbf{0}_k$. Let $\hat{\boldsymbol{\beta}}$ denote the solution.

(c) Show that $\|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2 = \|\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2$.

Problem 6. Suppose that V_1, V_2, V_3, and V_4 are independent standard normal random variables. Let $\bar{V} = \dfrac{V_1 + V_2 + V_3 + V_4}{4}$. Using the tables attached to this exam, compute (approximately) the following probabilities.

(a) $P\left(\bar{V} > 1 \text{ and } \sum_{i=1}^{4}(V_i - \bar{V})^2 < 1\right)$

(b) $P\left(\bar{V} > \sqrt{\displaystyle\sum_{i=1}^{4}(V_i - \bar{V})^2}\right)$

FORMULAS (for problems 7 and 8):

Suppose $\boldsymbol{y} \sim \text{Normal}_n(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n)$, $\boldsymbol{X}$ is a $n \times p$ full rank matrix, $n > p$, and $\boldsymbol{X}^\top \boldsymbol{X}$ is invertible. Let $\widehat{\boldsymbol{\beta}}$ be the MLE of $\boldsymbol{\beta}$ and $\widehat{\sigma}^2$ be the MLE of σ^2. Then $\dfrac{\|\boldsymbol{X}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\|^2/p}{\|\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2/(n-p)} \sim f_{p,n-p}$. If, in addition, $\boldsymbol{K}$ is a $m \times p$ full rank matrix (with $m \leq p$) such that $\boldsymbol{K}^\top \boldsymbol{\beta} = \boldsymbol{k}$, then

$$\begin{aligned} F &= \frac{(RSS_H - RSS)/m}{RSS/(n-p)} \\ &= \frac{\|\boldsymbol{X}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})\|^2/m}{\|\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2/(n-p)} \\ &= \frac{(\boldsymbol{K}^\top\widehat{\boldsymbol{\beta}} - \boldsymbol{k})^\top(\boldsymbol{K}^\top(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{K})^{-1}(\boldsymbol{K}^\top\widehat{\boldsymbol{\beta}} - \boldsymbol{k})/m}{(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})^\top(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}})/(n-p)} \sim f_{m,n-p} \end{aligned}$$

where $\widetilde{\boldsymbol{\beta}}$ is the restricted MLE of $\boldsymbol{\beta}$ satisfying $\boldsymbol{K}^\top\widetilde{\boldsymbol{\beta}} = \boldsymbol{k}$, $RSS_H = \|\boldsymbol{y} - \boldsymbol{X}\widetilde{\boldsymbol{\beta}}\|^2$, and $RSS = \|\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\|^2 = n\widehat{\sigma}^2$.

Problem 7. Suppose that $\boldsymbol{y}$ is a 4-dimensional Normal$(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I})$ random vector where $\boldsymbol{X}$ is a 4×2 matrix with columns $\boldsymbol{J}$ and $\boldsymbol{x}$, $\boldsymbol{J}$ is an 4-dimensional vector of ones, $\boldsymbol{x}$ is a non-random 4-dimensional vector, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ is a unknown vector of intercept and slope parameters, and σ^2 is the unknown variance parameter. Let $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\sigma}^2$ denote the maximum likelihood estimators of β_0, β_1 and σ^2, respectively.

(a) If $\boldsymbol{J}^T\boldsymbol{x} = 1$ and $\boldsymbol{x}^T\boldsymbol{x} = 3$, find values A, B, and C such that

$$A\left(\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}}\right)^2 + B\left(\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}}\right)\left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}}\right) + C\left(\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}}\right)^2 \leq 1$$

is a 95% confidence ellipse for $\boldsymbol{\beta}$.

(b) Now, also suppose that $\boldsymbol{J}^T\boldsymbol{y} = 2$, $\boldsymbol{y}^T\boldsymbol{y} = 38$, and $\boldsymbol{x}^T\boldsymbol{y} = 10$. Compute the maximum likelihood estimates $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\sigma}^2$. Is $\boldsymbol{\beta} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ inside the confidence ellipse in part (a)?

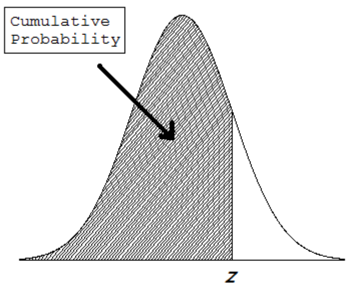
Problem 8. Suppose that

$$y_{ijk} = \mu_i + \alpha_{ij} + e_{ijk} \text{ for } i, j, k = 1, 2$$

where $\alpha_{11} + \alpha_{12} = \alpha_{21} + \alpha_{22} = 0$ and e_{ijk} are independent Normal$(0,\sigma^2)$ random variables.
Given the data

	y_{ijk}	$i = 1$	$i = 2$
j	1	7,9	6,4
	2	1,3	2,4

test $H_0 : \alpha_{11} = 0$ versus $H_0 : \alpha_{11} \neq 0$ at level 0.05. Make sure to state your conclusion and to show sufficient details (such as your test statistic and the corresponding critical value or bound for the P-value) to justify your conclusion.

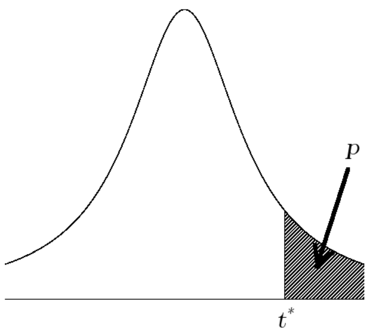


Cumulative probability for *z* is the area under the standard normal curve to the left of *z*.

Standard Normal Cumulative Probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

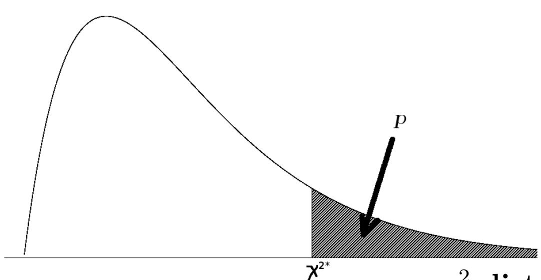
z	.00
3.5	0.999767
4.0	0.9999683
4.5	0.9999966
5.0	0.999999713



The critical value t^* is the value such that the area under the density curve of a t distribution with *df* degrees of freedom to the right of t^* is equal to p. It is also the value such that the area under the curve between $-t^*$ and t^* is equal to C.

t distribution critical values

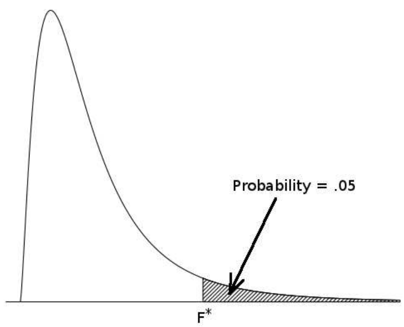
	Upper-tail probability p									
df	.05	.045	.04	.035	.03	.025	.02	.015	.01	.005
1	6.3	7.0	7.9	9.1	10.6	12.7	15.9	21.2	31.8	63.7
2	2.9	3.1	3.3	3.6	3.9	4.3	4.8	5.6	7.0	9.9
3	2.4	2.5	2.6	2.8	3.0	3.2	3.5	3.9	4.5	5.8
4	2.1	2.2	2.3	2.5	2.6	2.8	3.0	3.3	3.7	4.6
5	2.0	2.1	2.2	2.3	2.4	2.6	2.8	3.0	3.4	4.0
6	1.9	2.0	2.1	2.2	2.3	2.4	2.6	2.8	3.1	3.7
7	1.9	2.0	2.0	2.1	2.2	2.4	2.5	2.7	3.0	3.5
	90%	91%	92%	93%	94%	95%	96%	97%	98%	99%
	Confidence level C									



The critical value χ^{2*} is the value such that the area under the density curve of a χ^2 distribution with *df* degrees of freedom to the right of χ^{2*} is equal to p.

χ^2 distribution critical values

	Upper-tail probability p									
df	.90	.80	.70	.60	.50	.40	.30	.20	.10	.05
1	0.02	0.06	0.1	0.3	0.5	0.7	1.1	1.6	2.7	3.8
2	0.2	0.4	0.7	1.0	1.4	1.8	2.4	3.2	4.6	6.0
3	0.6	1.0	1.4	1.9	2.4	2.9	3.7	4.6	6.3	7.8
4	1.1	1.6	2.2	2.8	3.4	4.0	4.9	6.0	7.8	9.5
5	1.6	2.3	3.0	3.7	4.4	5.1	6.1	7.3	9.2	11.1
6	2.2	3.1	3.8	4.6	5.3	6.2	7.2	8.6	10.6	12.6
7	2.8	3.8	4.7	5.5	6.3	7.3	8.4	9.8	12.0	14.1



The critical value F^* is the value such that the area under the density curve of an F distribution with $df1$ degrees of freedom in the numerator and $df2$ degrees of freedom in the denominator to the right of F^* is equal to .05.

F distribution critical values

		df1 = degrees of freedom in the numerator										
		1	2	3	4	5	10	20	30	40	50	100
	1	161.45	199.50	215.71	224.58	230.16	241.88	248.01	250.10	251.14	251.77	253.04
	2	18.51	19.00	19.16	19.25	19.30	19.40	19.45	19.46	19.47	19.48	19.49
	3	10.13	9.55	9.28	9.12	9.01	8.79	8.66	8.62	8.59	8.58	8.55
	4	7.71	6.94	6.59	6.39	6.26	5.96	5.80	5.75	5.72	5.70	5.66
	5	6.61	5.79	5.41	5.19	5.05	4.74	4.56	4.50	4.46	4.44	4.41
	6	5.99	5.14	4.76	4.53	4.39	4.06	3.87	3.81	3.77	3.75	3.71
	7	5.59	4.74	4.35	4.12	3.97	3.64	3.44	3.38	3.34	3.32	3.27
	8	5.32	4.46	4.07	3.84	3.69	3.35	3.15	3.08	3.04	3.02	2.97
	9	5.12	4.26	3.86	3.63	3.48	3.14	2.94	2.86	2.83	2.80	2.76
	10	4.96	4.10	3.71	3.48	3.33	2.98	2.77	2.70	2.66	2.64	2.59
	11	4.84	3.98	3.59	3.36	3.20	2.85	2.65	2.57	2.53	2.51	2.46
	12	4.75	3.89	3.49	3.26	3.11	2.75	2.54	2.47	2.43	2.40	2.35
	13	4.67	3.81	3.41	3.18	3.03	2.67	2.46	2.38	2.34	2.31	2.26
	14	4.60	3.74	3.34	3.11	2.96	2.60	2.39	2.31	2.27	2.24	2.19
df2 =	15	4.54	3.68	3.29	3.06	2.90	2.54	2.33	2.25	2.20	2.18	2.12
degrees	16	4.49	3.63	3.24	3.01	2.85	2.49	2.28	2.19	2.15	2.12	2.07
of	17	4.45	3.59	3.20	2.96	2.81	2.45	2.23	2.15	2.10	2.08	2.02
freedom	18	4.41	3.55	3.16	2.93	2.77	2.41	2.19	2.11	2.06	2.04	1.98
in the	19	4.38	3.52	3.13	2.90	2.74	2.38	2.16	2.07	2.03	2.00	1.94
denominator	20	4.35	3.49	3.10	2.87	2.71	2.35	2.12	2.04	1.99	1.97	1.91
	21	4.32	3.47	3.07	2.84	2.68	2.32	2.10	2.01	1.96	1.94	1.88
	22	4.30	3.44	3.05	2.82	2.66	2.30	2.07	1.98	1.94	1.91	1.85
	23	4.28	3.42	3.03	2.80	2.64	2.27	2.05	1.96	1.91	1.88	1.82
	24	4.26	3.40	3.01	2.78	2.62	2.25	2.03	1.94	1.89	1.86	1.80
	25	4.24	3.39	2.99	2.76	2.60	2.24	2.01	1.92	1.87	1.84	1.78
	26	4.23	3.37	2.98	2.74	2.59	2.22	1.99	1.90	1.85	1.82	1.76
	27	4.21	3.35	2.96	2.73	2.57	2.20	1.97	1.88	1.84	1.81	1.74
	28	4.20	3.34	2.95	2.71	2.56	2.19	1.96	1.87	1.82	1.79	1.73
	29	4.18	3.33	2.93	2.70	2.55	2.18	1.94	1.85	1.81	1.77	1.71
	30	4.17	3.32	2.92	2.69	2.53	2.16	1.93	1.84	1.79	1.76	1.70
	40	4.08	3.23	2.84	2.61	2.45	2.08	1.84	1.74	1.69	1.66	1.59
	50	4.03	3.18	2.79	2.56	2.40	2.03	1.78	1.69	1.63	1.60	1.52
	100	3.94	3.09	2.70	2.46	2.31	1.93	1.68	1.57	1.52	1.48	1.39
	1000	3.85	3.00	2.61	2.38	2.22	1.84	1.58	1.47	1.41	1.36	1.26