Midterm Examination II[a]

STA 532: Statistical Inference

Friday, 2015 Apr 3, 8:30 – 9:45am

This is a closed-book exam. You may use a single sheet of prepared notes, if you wish, but you may not share materials.

If a question seems ambiguous or confusing, *please* ask me to clarify it. Unless a problem states otherwise, you must **show** your **work**. There are blank worksheets at the end of the test if you need more room for this (also normal & t tables and a list of pdfs etc. for common distributions).

It is to your advantage to write your solutions as clearly as possible, and to box answers I might not find.

For full credit, give answers in **closed form** (no unevaluated sums, integrals, maxima, unreduced fractions, *etc.*) where possible and **simplify**.

Good luck!

Print Name Clearly: ______________________________

1.	/20
2.	/20
3.	/20
4.	/20
5.	/20
Total:	/100

Problem 1: In each part of this problem the following statistics are from a random sample of n iid random variables $\{X_i\}$ from some specified probability distribution, but the specific distribution may vary:

$\frac{1}{n}\sum X_i = A$	$\frac{1}{n}\sum \log X_i = C$	$\max X_i = E$
$\frac{1}{n}\sum (X_i - \bar{X}_n)^2 = B$	$\frac{1}{n}\sum \log(1-X_i) = D$	$\min X_i = F$

In each case, find a statistic T that is a function of one or more of n, A, B, C, D, E, F for which the Rejection Region of the indicated LRT or GLRT is of the form $\mathcal{R} = \{x : T(x) \ge t\}$. For four points each (two for picking the right statistics among $ABCDEF$, two for **a** right function of them for T):

a) $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{Ex}(\lambda)$; $H_0 : \lambda = 1$ vs. $H_1 : \lambda = 2$.

$T =$

b) $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{No}(\mu, 1)$; $H_0 : \mu = 1$ vs. $H_1 : \mu = 2$.

$T =$

c) $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{No}(\mu, \sigma^2)$; $H_0 : \mu = 1$ vs. $H_1 : \mu \neq 1$ (σ^2 unknown).

$T =$

d) $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{No}(3, \sigma^2)$; $H_0 : \sigma^2 = 9$ vs. $H_1 : \sigma^2 = 25$.

$T =$

e) $H_0 : \{X_i\} \overset{\text{iid}}{\sim} \mathsf{Ex}(\lambda)$ vs. $H_1 : \{X_i\} \overset{\text{iid}}{\sim} \mathsf{Un}(0, \theta)$.

$T =$

Problem 2: If $\{X_i\}_{1\le i\le 10} \overset{\text{iid}}{\sim} \mathsf{No}(\mu,\sigma^2)$ with σ^2 unknown but μ known, then three different valid 90% confidence intervals for σ^2 are:

a. Since $W = n(\bar{X}-\mu)^2$ has $\frac{W}{\sigma^2} \sim \chi^2_1$, and $0.90 = \mathsf{P}[0.00393 \le \chi^2_1 < 3.841]$,

$$[W/3.841 \le \sigma^2 \le W/0.00393]$$

b. Since $S = \sum(X_i - \bar{X}_n)^2$ has $\frac{S}{\sigma^2} \sim \chi^2_9$ and $0.90 = \mathsf{P}[3.325 \le \chi^2_9 \le 16.92]$,

$$\big[S/16.92 \le \sigma^2 \le S/3.325\big]$$

c. Since $Y = \sum(X_i - \mu)^2$ has $\frac{Y}{\sigma^2} \sim \chi^2_{10}$ and $0.90 = \mathsf{P}[3.940 \le \chi^2_{10} \le 18.31]$,

$$\big[Y/18.31 \le \sigma^2 \le Y/3.940\big].$$

Which of these is best, and why? Or are they all equally good?

Problem 3: Eric and Erica are testing hypotheses about independent Poisson random variables $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{Po}(\lambda)$ and $\{Y_j\} \overset{\text{iid}}{\sim} \mathsf{Po}(\theta)$.

a) (5) Eric chose a sample size $n \in \mathbb{N}$ and took n observations. When he found that all were zero, $X_1 = \cdots = X_n = 0$, he announced that the hypothesis $H_0 : \lambda = 2$ was rejected (in favor of $H_1 : \lambda = 1$) at level $\alpha = 0.0001 = 10^{-4}$. How large was n?

b) (5) Erica chose a sample size $m \in \mathbb{N}$ and took m observations. When she found that all were zero, $Y_1 = \cdots = Y_m = 0$, she announced that the posterior probability of the hypothesis $H_0 : \theta = 2$ (as opposed to $H_1 : \theta = 1$) was below $0.0001 = 10^{-4}$. She began with equal prior probabilities $\pi_0 = \mathsf{P}[H_0] = 1/2$ and $\pi_1 = \mathsf{P}[H_1] = 1/2$. How large was m?

Problem 3 (cont'd): With samples $\{X_1, \ldots, X_{10}\} \stackrel{\text{iid}}{\sim} \mathsf{Po}(\lambda)$ and $\{Y_1, \ldots, Y_{10}\} \stackrel{\text{iid}}{\sim} \mathsf{Po}(\theta)$ (both of size 10) from two new Poisson populations (so you can ignore the observations from parts a) and b) above),

c) (10) Eric and Erica agree to test the hypothesis $H_0 : \lambda = \theta$ against the two-sided alternative $H_1 : \lambda \neq \theta$ using the GLRT. Find the Generalized Likelihood Ratio statistic they will need to test H_0, in terms of the sufficient statistics $S = \sum_{i=1}^{10} X_i$ and $T = \sum_{j=1}^{10} Y_j$. Simplify.
$\Lambda =$

Problem 4: Inter-arrival times for buses, tellers, synapse spikes, etc. are sometimes modeled with exponential distributions and sometimes with gamma distributions with smaller variability than exponentials. Of course the exponential is a special case of the gamma, so let's consider a test of

$$H_0 : \alpha = 1 \qquad \text{vs.} \qquad H_1 : \alpha = 2$$

for data $\{X_i : \ 1 \le i \le n\} \overset{\text{iid}}{\sim} \mathsf{Ga}(\alpha, \lambda)$, with unknown rate λ. Give all answers as simply as possible **in terms of** one or more of these **summary statistics**:

$$S = \frac{1}{n}\sum X_i \qquad T = \frac{1}{n}\sum \log X_i \qquad U = \frac{1}{n}\sum (X_i - \bar{X}_n)^2.$$

a) (5) Show that the conditional MLE of λ for any specified α is $\hat{\lambda}(\alpha) = \alpha/S$

b) (5) Evaluate the likelihood function $f(\mathbf{x} \mid \alpha, \lambda)$ for the data $\mathbf{x} = \{X_i\}$ at the conditional MLE $\lambda = \hat{\lambda}(\alpha)$ as a function of α, n, and S, T, U:
$f\big(\mathbf{x} \mid \alpha, \hat{\lambda}(\alpha)\big) =$

Problem 4 (cont'd): Still $\{X_i\} \overset{\text{iid}}{\sim} \mathsf{Ga}(\alpha, \lambda)$ and $S = \frac{1}{n}\sum X_i$, $T = \frac{1}{n}\sum \log X_i$, and $U = \frac{1}{n}\sum(X_i - \bar{X}_n)^2$:

c) (10) Find the GLR statistic for $H_1 : \alpha = 2$ against $H_0 : \alpha = 1$ in terms of n, S, T, U:
$\Lambda =$

Problem 5: True or false? No explanations needed, unless you find a problem ambiguous.

a. T F The Beta prior distribution is conjugate for the Bernoulli sampling distribution.

b. T F If X has a χ^2_ν distribution then $\mathsf{E}X = \nu$.

c. T F The estimator with the smallest MSE is always unbiased.

d. T F If $\{X_i\} \sim \mathsf{Ga}(\alpha, 2)$ then $\bar{X}$ is sufficient for α.

e. T F UMP tests exist for all one-sided hypotheses.

f. T F Reject H_0 if you observe data X in the critical region $\mathcal{R}$.

g. T F The *power* function $[1 - \beta(\theta)]$ is the probability that H_0 is rejected, for $\theta \in H_1$.

h. T F If the P-value is below 0.05, then $\mathsf{P}[H_0] \le 5\%$.

i. T F If the P-value is above 0.99, *reject* at level $\alpha = 0.01$.

j. T F One way to test $H_0 : \theta = \theta_0$ at size $\alpha = 0.10$ is to construct a 90% confidence set $C(X)$ and reject if $\theta_0 \notin C(X)$.

Extra worksheet, if needed:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz:$$

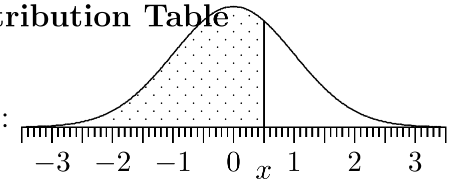


Table 5.1 Area $\Phi(x)$ under the Standard Normal Curve to the left of x.

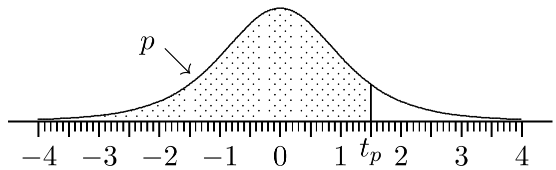
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.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

$\Phi(0.6745) = 0.75$ $\Phi(1.6449) = 0.95$ $\Phi(2.3263) = 0.99$ $\Phi(3.0902) = 0.999$

$\Phi(1.2816) = 0.90$ $\Phi(1.9600) = 0.975$ $\Phi(2.5758) = 0.995$ $\Phi(3.2905) = 0.9995$

Critical Values for Student's t

$$p = \int_{-\infty}^{t_p} c \, \frac{dt}{(1 + t^2/\nu)^{(\nu+1)/2}}$$



ν	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$	$t_{.9999}$
1	0.325	0.727	1.376	1.9626	3.078	6.314	12.76	31.82	63.66	318.3	636.6	3183.
2	0.289	0.617	1.061	1.3862	1.886	2.920	4.303	6.965	9.925	22.33	31.60	70.70
3	0.277	0.584	0.978	1.2498	1.638	2.353	3.182	4.541	5.841	10.22	12.92	22.20
4	0.271	0.569	0.941	1.1896	1.533	2.132	2.776	3.747	4.604	7.173	8.610	13.03
5	0.267	0.559	0.920	1.1558	1.476	2.015	2.571	3.365	4.032	5.893	6.869	9.678
6	0.265	0.553	0.906	1.1342	1.440	1.943	2.447	3.143	3.707	5.208	5.959	8.025
7	0.263	0.549	0.896	1.1192	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7.063
8	0.262	0.546	0.889	1.1081	1.397	1.860	2.306	2.896	3.355	4.501	5.041	6.442
9	0.261	0.543	0.883	1.0997	1.383	1.833	2.262	2.821	3.250	4.297	4.781	6.010
10	0.260	0.542	0.879	1.0931	1.372	1.812	2.228	2.764	3.169	4.144	4.587	5.694
11	0.260	0.540	0.876	1.0877	1.363	1.796	2.201	2.718	3.106	4.025	4.437	5.453
12	0.259	0.539	0.873	1.0832	1.356	1.782	2.179	2.681	3.055	3.930	4.318	5.263
13	0.259	0.538	0.870	1.0795	1.350	1.771	2.160	2.650	3.012	3.852	4.221	5.111
14	0.258	0.537	0.868	1.0763	1.345	1.761	2.145	2.624	2.977	3.787	4.140	4.985
15	0.258	0.536	0.866	1.0735	1.341	1.753	2.131	2.602	2.947	3.733	4.073	4.880
16	0.258	0.535	0.865	1.0711	1.337	1.746	2.120	2.583	2.921	3.686	4.015	4.791
17	0.257	0.534	0.863	1.0690	1.333	1.740	2.110	2.567	2.898	3.646	3.965	4.714
18	0.257	0.534	0.862	1.0672	1.330	1.734	2.101	2.552	2.878	3.610	3.922	4.648
19	0.257	0.533	0.861	1.0655	1.328	1.729	2.093	2.539	2.861	3.579	3.883	4.590
20	0.257	0.533	0.860	1.0640	1.325	1.725	2.086	2.528	2.845	3.552	3.85	4.539
21	0.257	0.532	0.859	1.0627	1.323	1.721	2.080	2.518	2.831	3.527	3.819	4.493
22	0.256	0.532	0.858	1.0614	1.321	1.717	2.074	2.508	2.819	3.505	3.792	4.452
23	0.256	0.532	0.858	1.0603	1.319	1.714	2.069	2.500	2.807	3.485	3.768	4.415
24	0.256	0.531	0.857	1.0593	1.318	1.711	2.064	2.492	2.797	3.467	3.745	4.382
25	0.256	0.531	0.856	1.0584	1.316	1.708	2.060	2.485	2.787	3.450	3.725	4.352
26	0.256	0.531	0.856	1.0575	1.315	1.706	2.056	2.479	2.779	3.435	3.707	4.324
27	0.256	0.531	0.855	1.0567	1.314	1.703	2.052	2.473	2.771	3.421	3.690	4.299
28	0.256	0.530	0.855	1.0560	1.313	1.701	2.048	2.467	2.763	3.408	3.674	4.275
29	0.256	0.530	0.854	1.0553	1.311	1.699	2.045	2.462	2.756	3.396	3.659	4.254
30	0.256	0.530	0.854	1.0547	1.310	1.697	2.042	2.457	2.750	3.385	3.646	4.234
40	0.255	0.529	0.851	1.0500	1.303	1.684	2.021	2.423	2.704	3.307	3.551	4.094
60	0.254	0.527	0.848	1.0455	1.296	1.671	2.000	2.390	2.660	3.232	3.460	3.962
120	0.254	0.526	0.845	1.0409	1.289	1.658	1.980	2.358	2.617	3.160	3.373	3.837
∞	0.253	0.524	0.842	1.0364	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.719

Name	Notation	pdf/pmf	Range	Mean μ	Variance σ^2	
Beta	$\mathsf{Be}(\alpha,\beta)$	$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$x\in(0,1)$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	
Binomial	$\mathsf{Bi}(n,p)$	$f(x)=\binom{n}{x}p^x q^{(n-x)}$	$x\in 0,\cdots,n$	$n\,p$	$n\,p\,q$	$(q=1-p)$
Exponential	$\mathsf{Ex}(\lambda)$	$f(x)=\lambda\,e^{-\lambda x}$	$x\in\mathbb{R}_+$	$1/\lambda$	$1/\lambda^2$	
Gamma	$\mathsf{Ga}(\alpha,\lambda)$	$f(x)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\,e^{-\lambda x}$	$x\in\mathbb{R}_+$	α/λ	α/λ^2	
Geometric	$\mathsf{Ge}(p)$	$f(x)=p\,q^x$	$x\in\mathbb{Z}_+$	q/p	q/p^2	$(q=1-p)$
		$f(y)=p\,q^{y-1}$	$y\in\{1,...\}$	$1/p$	q/p^2	$(y=x+1)$
HyperGeo.	$\mathsf{HG}(n,A,B)$	$f(x)=\frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}$	$x\in 0,\cdots,n$	$n\,P$	$n\,P\,(1-P)\frac{N-n}{N-1}$	$(P=\frac{A}{A+B})$
Logistic	$\mathsf{Lo}(\mu,\beta)$	$f(x)=\frac{e^{-(x-\mu)/\beta}}{\beta[1+e^{-(x-\mu)/\beta}]^2}$	$x\in\mathbb{R}$	μ	$\pi^2\beta^2/3$	
Log Normal	$\mathsf{LN}(\mu,\sigma^2)$	$f(x)=\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-(\log x-\mu)^2/2\sigma^2}$	$x\in\mathbb{R}_+$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$	
Neg. Binom.	$\mathsf{NB}(\alpha,p)$	$f(x)=\binom{x+\alpha-1}{x}p^\alpha\,q^x$	$x\in\mathbb{Z}_+$	$\alpha q/p$	$\alpha q/p^2$	$(q=1-p)$
		$f(y)=\binom{y-1}{y-\alpha}p^\alpha\,q^{y-\alpha}$	$y\in\{\alpha,...\}$	α/p	$\alpha q/p^2$	$(y=x+\alpha)$
Normal	$\mathsf{No}(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$	$x\in\mathbb{R}$	μ	σ^2	
Pareto	$\mathsf{Pa}(\alpha,\epsilon)$	$f(x)=\alpha\,\epsilon^\alpha/(x+\epsilon)^{\alpha+1}$	$x\in\mathbb{R}_+$	$\frac{\epsilon}{\alpha-1}$	$\frac{\epsilon^2\alpha}{(\alpha-1)^2(\alpha-2)}$	
		$f(y)=\alpha\,\epsilon^\alpha/y^{\alpha+1}$	$y\in(\epsilon,\infty)$	$\frac{\epsilon\,\alpha}{\alpha-1}$	$\frac{\epsilon^2\alpha}{(\alpha-1)^2(\alpha-2)}$	$(y=x+\epsilon)$
Poisson	$\mathsf{Po}(\lambda)$	$f(x)=\frac{\lambda^x}{x!}e^{-\lambda}$	$x\in\mathbb{Z}_+$	λ	λ	
Snedecor F	$F(\nu_1,\nu_2)$	$f(x)=\frac{\Gamma(\frac{\nu_1+\nu_2}{2})(\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\times$ $x^{\frac{\nu_1-2}{2}}\left[1+\frac{\nu_1}{\nu_2}\,x\right]^{-\frac{\nu_1+\nu_2}{2}}$	$x\in\mathbb{R}_+$	$\frac{\nu_2}{\nu_2-2}$	$\left(\frac{\nu_2}{\nu_2-2}\right)^2\frac{2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-4)}$	
Student t	$t(\nu)$	$f(x)=\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}[1+x^2/\nu]^{-(\nu+1)/2}$	$x\in\mathbb{R}$	0	$\nu/(\nu-2)$	
Uniform	$\mathsf{Un}(a,b)$	$f(x)=\frac{1}{b-a}$	$x\in(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	
Weibull	$\mathsf{We}(\alpha,\beta)$	$f(x)=\alpha\beta\,x^{\alpha-1}\,e^{-\beta\,x^\alpha}$	$x\in\mathbb{R}_+$	$\frac{\Gamma(1+\alpha^{-1})}{\beta^{1/\alpha}}$	$\frac{\Gamma(1+2/\alpha)-\Gamma^2(1+1/\alpha)}{\beta^{2/\alpha}}$	