Midterm Examination

Mth 136 = Sta 114

Wednesday, 2001 February 21, 2:20 – 3:35 pm

This is a closed-book examination so please do not refer to your notes, the text, or to any other books. You may use a one-sided single sheet of *your own* notes, if you wish, but you may not share materials. A normal distribution table, the PDF handout, and a blank worksheet are attached to the exam. If you don't understand something in one of the questions feel free to ask me, but please do not talk to each other. You may use a calculator but not a laptop computer.

You must **show** your **work** to get partial credit. Unsupported answers are not acceptable, even if they are correct. Please give all numerical answers as fractions **in lowest terms** (simplify!) or as decimals correct to **four places**. You should spend about 10–15 minutes on each problem. It is to your advantage to write your solutions as clearly as possible, since I cannot give credit for solutions I do not understand. Good luck.

Cheating on exams is a breach of trust with classmates and faculty, and will not be tolerated. After completing the exam please acknowledge the Duke Honor Code:

I have neither given nor received any unauthorized aid on this exam.

Signature: ______________________

Print Name: ______________________

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

Problem 1: A single blue marble is in a bag with some number θ of white marbles, with θ known to be one of the elements of $\Theta = \{0, 1, 2, 3\}$, but otherwise unknown. Repeatedly we perform the experiment of drawing a single marble at random from the bag; noting its color; then returning it to the bag. Denote by B the number of Blue marbles discovered in N independent repetitions of the experiment.

In $N = 4$ repetitions of this experiment we find a blue marble three times and a white marble once.

a) (5) Find the likelihood function $f_4(B = 3 \mid \theta)$ for this data.

b) (5) Find the maximum likelihood estimator $\hat{\theta} \in \Theta$ for $B = 3$.

c) (10) Beginning with a uniform prior distribution that places probability $\xi(\theta) = 0.25$ on each of the four possible values of $\theta \in \Theta$, give the posterior distribution (*i.e.*, give $\xi(\theta|B = 3)$, $\theta \in \Theta$).

Problem 2: The times-to-failure T_j of an electronic component called quixars are independent random variables with pdf's

$$f(x|\lambda) = \lambda^2 x e^{-\lambda x}, \qquad x > 0$$

for some unknown $\lambda > 0$, for $1 \leq j \leq n$. We have $n = 4$ observations, with failure-times $T_1 = 2$, $T_2 = 3$, $T_3 = 5$, $T_4 = 6$.

a) (5) For these data, find the Likelihood Function $f_n(x \mid \lambda)$;

b) (5) Find the maximum likelihood estimate $\hat{\lambda}_n(x)$:

c) (5) With a uniform prior $\xi(\lambda) \equiv 1$, find the posterior density function $\xi(\lambda \mid x)$ (or the name of the posterior distribution, with the value(s) of any parameter(s)):

d) (5) Give the posterior mean $\mathsf{E}[\lambda \mid x]$ for this prior (Note: remember a table of familiar pdf's is attached to this test).

Problem 3: Twenty-five measurements X_j come from a normal distribution with mean $\mathsf{E}[X_j] = \mu$ and variance $\mathsf{E}[(X_j - \mu)^2] = 9$. Find the Maximum Likelihood Estimator and a (frequentist) 90% Confidence Interval for μ, *i.e.*, an interval $[L(x), R(x)]$ with $\mathsf{P}[L(x) < \mu < R(x)] = 0.90$. Give the endpoints for data that satisfy:

$$\sum_{j=1}^{25} X_j = 250 \qquad\qquad \sum_{j=1}^{25} X_j^2 = 3725$$

a) (10) Find the Maximum Likelihood Estimator for μ (you don't have to prove it's the MLE, just show how you found its value):
$\hat{\mu} =$ ______________

b) (10) Find the Confidence Interval:
L(x)=______________ R(x)=______________

Warning: Sometimes you have more information than you need...

Problem 4: Stacy and Toby each discovered a possible cure for cancer; let θ_S and θ_T be the probability of cure for a cancer subject under Stacy's and Toby's treatments, respectively. In a randomized clinical trial (RCT), $N_S = 25$ subjects are given Stacy's treatment (of whom $S = 10$ show marked improvement) and $N_T = 30$ subjects are given Toby's treatment, of whom $T = 20$ show marked improvement.

Our statistician Chris elects to use independent uniform prior distributions $\xi(\theta_S) \equiv 1,\ 0 < \theta_S < 1$ and $\xi(\theta_T) \equiv 1,\ 0 < \theta_T < 1$.

a) (8) What are the posterior distributions for θ_S and θ_T? Give the pdfs **or** (better) the name and value(s) of any parameter(s).
$S \sim$ ____________ $T \sim$ ____________

b) (8) What are the posterior means and variances for the two treatments? Remember the attached pdf sheet...
$\mathsf{E}[\theta_S \mid S = 10] =$ ________ $\mathsf{Var}[\theta_S \mid S = 10] =$ ________
$\mathsf{E}[\theta_T \mid T = 20] =$ ________ $\mathsf{Var}[\theta_T \mid T = 20] =$ ________

c) (4) If the posterior distributions of θ_S and θ_T are well enough approximated by normal distributions, what would be the approximate probability $\mathsf{P}[\theta_S > \theta_T]$? Why? (Hint: These are the hardest 4 points on this test)
$\mathsf{P}[\theta_S > \theta_T] \approx$ ________________________

Problem 5: Jamie and Corey are trying to find interval estimates for the mean μ of a normal distribution with unknown variance σ^2, on the basis of 9 independent observations $X_1, ..., X_9$. They use Maximum Likelihood Estimation to find estimates

$$\hat{\mu}_9 = \bar{X}_9 = \frac{1}{9}\sum X_j = 55.0 \qquad \hat{\sigma}_9^2 = S_9^2 = \frac{1}{9}\sum(X_j - \bar{X}_9)^2 = 2048.0$$

Jamie uses a Normal Distribution (Z) table to find a 90% interval of the form $\mu = \bar{X}_n \pm 1.645\sqrt{S_n^2/n}$, while Corey uses the Student's t distribution with $n-1 = 8$ degrees of freedom to find an interval.

a) (4) Whose interval will be *shorter*? Why?

b) (4) What is the approximate probability that an interval formed using Jamie's method will include μ? Choose one and explain **briefly**:
○ Less than 90% ○ About 90% ○ More than 90%?

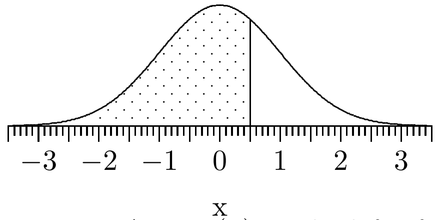
c) (4) What is the approximate probability that an interval formed using Corey method will include μ? Choose one and explain:
○ Less than 90% ○ About 90% ○ More than 90%?

d) (4) Which of the following is the approximate distribution of S_9^2, the MLE for σ^2? Mark all that are correct (if any are). The subscripts on t and χ^2 indicate degrees of freedom. No explanations needed.
○ $\mathsf{No}(\mu, \sigma^2)$ ○ $\mathsf{Ga}(4, \frac{4.5}{\sigma^2})$ ○ $\mathsf{Be}(0.55, 0.45)$ ○ t_8 ○ $\chi_8^2 \times \frac{\sigma^2}{9}$

e) (4) Who constructed the best 90% confidence interval? Why?
○ Jamie ○ Corey ○ Both equally good

Extra worksheet, if needed:

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



Normal Distribution Table Area $\Phi(x)$ 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

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

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