Name______________________________

IE 4521 – Midterm #2

Prof. John Gunnar Carlsson

April 12, 2011

Before you begin: This exam has 10 pages (including the quantile tables for the t- and F-distributions) and a total of 6 problems. Make sure that all pages are present. To obtain credit for a problem, you must show all your work. if you use a formula to answer a problem, write the formula down. Do not open this exam until instructed to do so.

1. (10 points) Let X be a random variable with the following probability distribution:

$$f(x) = \begin{cases} (\theta+1)\, x^{\theta} & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Use the method of moments to estimate θ, given samples $x_1, \ldots, x_n$.
   We have $E(X) = \int_0^1 x f(x)\, dx = \int_0^1 (\theta+1)\, x^{\theta+1}\, dx = \frac{\theta+1}{\theta+2}$, i.e. $\theta = \frac{1-2E(X)}{E(X)-1}$, and thus $\hat{\theta} = \frac{1-2\bar{x}}{\bar{x}-1}$.

   (b) Find the maximum likelihood estimator of θ, given a collection of samples $x_1, \ldots, x_n$.
   We have

$$\begin{aligned} \mathcal{L}(\theta) &= \left[(\theta+1)\, x_1^{\theta}\right] \times \cdots \times \left[(\theta+1)\, x_n^{\theta}\right] \\ &= (\theta+1)^n (x_1 \cdots x_n)^{\theta} \\ \ell(\theta) &= n \log(\theta+1) + \theta \sum_i \log(x_i) \\ \frac{d\ell}{d\theta} &= \frac{n}{\theta+1} + \sum_i \log(x_i) = 0 \\ \hat{\theta} &= \frac{-n}{\sum_i \log(x_i)} - 1 \end{aligned}$$

2. (20 points)

(a) Consider the Poisson distribution with parameter μ:

$$f(x) = \frac{\mu^x}{x!} \cdot e^{-\mu}$$

Find the maximum likelihood estimator of μ, given a collection of samples $x_1, \ldots, x_n$.
We have

$$\begin{aligned}
\mathcal{L}(\mu) &= \left(\frac{\mu^{x_1}}{x_1!} \cdot e^{-\mu}\right) \times \cdots \times \left(\frac{\mu^{x_n}}{x_n!} \cdot e^{-\mu}\right) \\
&= \frac{\mu^{x_1 + \cdots x_n}}{x_1! \cdots x_n!} e^{-\mu n} \\
\ell(\mu) &= (x_1 + \cdots + x_n) \log \mu - \log(x_1! \cdots x_n!) - \mu n \\
\frac{d\ell}{d\mu} &= \frac{x_1 + \cdots + x_n}{\mu} - n = 0 \\
\hat{\mu} &= \frac{x_1 + \cdots + x_n}{n}
\end{aligned}$$

(b) Consider the *shifted exponential distribution*:

$$f(x) = \lambda e^{-\lambda(x-\theta)},\ x \geq \theta$$

i. Write the expressions for $\partial\ell/\partial\lambda$ and $\partial\ell/\partial\theta$, where $\ell(\cdot)$ is the log-likelihood function.
We have

$$\begin{aligned}
\mathcal{L}(\lambda, \theta) &= \left[\lambda e^{-\lambda(x_1-\theta)}\right] \times \cdots \times \left[\lambda e^{-\lambda(x_n-\theta)}\right] \\
&= \lambda^n e^{-\lambda(x_1+\cdots+x_n-n\theta)} \\
\ell(\lambda, \theta) &= n \log \lambda - \lambda(x_1 + \cdots + x_n - n\theta) \\
\frac{\partial\ell}{\partial\lambda} &= \frac{n}{\lambda} + n\theta - (x_1 + \cdots + x_n) \\
\frac{\partial\ell}{\partial\theta} &= \lambda n
\end{aligned}$$

ii. Using the previous result, find the maximum likelihood estimators of λ and θ, given a collection of samples $x_1, \ldots, x_n$. (*Hint: you may find that $\partial\ell/\partial\theta$ is always positive for $\lambda > 0$; this suggests that $\ell(\cdot)$ is always increasing as a function of θ. However, recall that we must have $x_i \geq \theta$ for all samples x_i. What does this tell you about $\hat{\theta}$?*).
We should set $\hat{\theta}$ to be as large as possible, so that $\hat{\theta} = \min_i\{x_i\}$. This lets us solve for $\hat{\lambda} = \frac{n}{x_1+\cdots+x_n-n\hat{\theta}} = \frac{n}{x_1+\cdots+x_n-n\min_i\{x_i\}}$.

(c) (Bonus; 2 points) Describe a practical situation in which one would suspect that the shifted exponential distribution is a plausible model.

3. (20 points) In a 1943 experiment (Whitlock and Bliss, “A Bioassay Technique for Antihelminthics,” *Journal of Parasitology*, **29**, pp. 48–58), 10 albino rats were used to study the effectiveness of carbon tetrachloride as a treatment for worms. Each rat received an injection of worm larvae. After 8 days, the rats were randomly divided into two groups of 5 each; each rat in the first group received a dose of 0.032 cc of carbon tetrachloride, whereas the dosage for each rat in the second group was 0.063 cc. Two days later the rats were killed, and the number of adult worms in each rat was determined. The numbers detected in the group receiving the .032 dosage were

$$421, 462, 400, 378, 413$$

whereas they were

$$207, 17, 412, 74, 116$$

for those receiving the .063 dosage. Using the method of paired samples, perform *two* one-sided hypothesis tests (one in each direction) to determine if the larger dosage is more effective than the smaller, and determine what conclusion you can make from them. Use $\alpha = 0.05$.

We'll perform a paired-sample test with $n = 5$. The new samples of our variable Z are $214, 445, -12, 304, 297$. We're testing the null hypotheses $H_0 : \mu \geq 0$ and $H_0 : \mu \leq 0$, where $\mu = E(Z)$. We find that $s = 168.$ and $\bar{z} = 250.$, so the t-statistic is

$$t_0 = \frac{\sqrt{n}\,(\bar{z} - \mu_0)}{s} = \frac{\sqrt{5}\,(250. - 0)}{168} \approx 3.33.$$

Looking in the table, we see that $t_{\alpha,n-1} = t_{0.05,4} = 2.132$. Since $t_0 > 2.132$ and $t_0 \geq -2.132$, we should reject the null hypothesis that $\mu \leq 0$ and accept the null hypothesis that $\mu \geq 0$, and therefore we conclude that it is implausible (at the 95% significance level) that the smaller dosage is as effective as the larger dosage.

This page intentionally left blank.

4. (10 points) The following data summary was obtained from a comparison of the lead content of human hair removed from adult individuals that had died between 1880 and 1920 with the lead content of present-day adults. The data are in units of micrograms, equal to one-millionth of a gram.

	1880-1920	**Today**
Sample size	30	100
Sample mean	48.5	26.6
Sample standard deviation	14.5	12.3

Perform a two-sided hypothesis tests and determine what conclusion you can make from them. Use $\alpha = 0.05$.

We'll use the general method to compare two populations. We'll test the null hypotheses $H_0 : \mu_A - \mu_B \geq 0$ and $H_0 : \mu_A - \mu_B \leq 0$. The t-statistic is given by

$$t_0 = \frac{\bar{x} - \bar{y} - \delta}{\sqrt{\frac{S_x^2}{n} + \frac{S_y^2}{m}}} = \frac{48.5 - 26.6 - 0}{\sqrt{\frac{14.5^2}{30} + \frac{12.3^2}{100}}} \approx 7.50.$$

We'll use $\nu = \min\{n, m\} - 1 = 29$, and so looking in the table we find that $t_{\alpha/2,\nu} = t_{0.025,29} = 2.045$. Since $t_0 > 2.045$, we should reject the null hypothesis that $\mu_A - \mu_B = 0$ and therefore we conclude that it is implausible (at the 95% significance level) that the average lead content has remained the same between 1880 and 1920.

5. (15 points) Four standard chemical procedures are used to determine the magnesium content in a certain chemical compound. Each procedure is used four times on a given compound with the following data resulting:

	Magnesium content			
	Method 1	Method 2	Method 3	Method 4
	76.42	80.41	74.20	86.20
	78.62	82.26	72.68	86.04
	80.40	81.15	78.84	84.36
	78.20	79.20	80.32	80.68
Population means	78.41	80.76	76.51	84.32

The mean of all samples is $\bar{x}_{\cdot\cdot} = 80.00$. The differences $x_{ij} - \bar{x}_{i\cdot}$ are given in the table below:

Deviations from $\bar{x}_{i\cdot}$			
Method 1	Method 2	Method 3	Method 4
-1.99	-0.34	-2.31	1.88
0.21	1.50	-3.83	1.72
1.99	0.40	2.33	0.04
-0.21	-1.56	3.81	-3.64

Do the data indicate, at the $\alpha = 0.1$ significance level, that the procedures yield equivalent results?

We have

$$\begin{aligned}
\text{SSTr} &= \sum_{i=1}^{k} n_i \left(\bar{x}_{i\cdot} - \bar{x}_{\cdot\cdot}\right)^2 = 135. \\
\text{MSTr} &= \frac{\text{SSTr}}{k-1} = 45.0 \\
\text{SSE} &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_{i\cdot}\right)^2 = 72.7 \\
\text{MSE} &= \frac{\text{SSE}}{n_T - k} = 6.06
\end{aligned}$$

and therefore $F_0 = \frac{\text{MSTr}}{\text{MSE}} = 7.42$. Looking up in the table, we see that $F_{k-1,n_T-k,\alpha} = F_{3,12,0.1} = 2.61$, and therefore we should accept the null hypothesis.

6. (10 points) True/false. No justification is needed.

(a) Other things being equal, as the confidence level $1 - \alpha$ for a confidence interval increases, the width of the interval increases. True.

(b) The method of moments can sometimes return parameter estimates for which the samples $x_1, \ldots, x_n$ are infeasible. True.

(c) The maximum likelihood estimate of σ^2 in a normal distribution is biased. True.

(d) A p-value is the probability of obtaining the given data set or worse when the null hypothesis H_0 is true. True.

(e) For an exponential distribution, both the method of moments and maximum likelihood estimation give $\hat{\lambda} = 1/\bar{x}$. True.

(f) Maximum likelihood estimation only applies to continuous random variables, since a p.m.f. cannot be differentiated with respect to x. False.

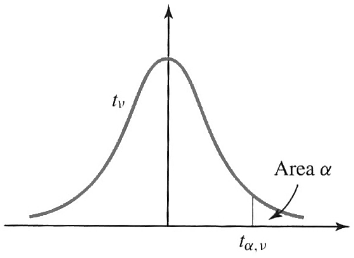
(g) A biased estimator always has a greater mean square error than an unbiased estimator. False.

(h) If X is uniformly distributed between 0 andθ, then both the method of moments and the maximum likelihood estimators for θ are biased. False.

(i) In the construction of confidence intervals, if all other quantities are unchanged, an increase in the sample size will lead to a wider interval. False.

(j) The p-value for a hypothesis test is inversely proportional to the actual probability that the null hypothesis is true. False.

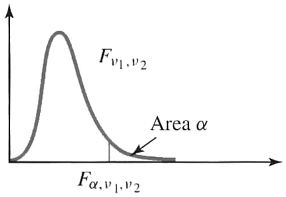
Table III: Critical Points of the t-Distribution



Degrees of freedom ν				α			
	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table IV: Critical Points of the F-Distribution

$$F_{\nu_1,\nu_2} \sim \frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$$



$\alpha = 0.10$

ν_1

ν_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	∞
1	39.86	49.50	53.59	55.84	57.24	58.20	58.90	59.44	59.85	60.20	60.70	61.22	61.74	62.00	62.27	62.53	62.79	63.05	63.33
2	3.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.28	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.51
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.35
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.18
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.10