//

STATISTICS DEPARTMENT
M.S. EXAMINATION

PART I
CLOSED BOOK

Friday, May 14, 2004

9:00 a.m. - 1:00 p.m.

Biella Room (Library, First Floor)

Instructions: **Complete *all four* problems. Each problem counts 25 points. Unless otherwise noted, points are allocated approximately equally to lettered parts of a problem. Spend your time accordingly.**

Begin each problem on a new page. Write the problem number and the page number in the specified locations at the top of each page. Also write your chosen ID code number on every page. Please write only within the black borderlines, leaving at least 1" margins on both sides, top and bottom of each page. Write on one side of the page only.

At the end of this part of the exam you will turn in your answers sheets, but you will keep the question sheets and your scratch paper.

Tables of some distributions are provided. Use them as appropriate.

1. Cars enter a certain stretch of the freeway at an average rate of $1/\beta$ per minute, according to a Poisson distribution. Starting from a fixed point of time, let Y be the time (in minutes) until the entrance of the α^{th} car ($\alpha = 1, 2, 3, \ldots$) into this stretch of freeway. It can be shown that

$$P(Y > y) = P(X \leq \alpha - 1), \qquad (*)$$

where X is a Poisson random variable with mean $= \lambda = y/\beta$. In working the parts below, suppose $\beta = .1$.

a) Let Y_1 be the time (in minutes) until the entrance of the 1^{st} car into this stretch of freeway. Write down the probability density function $f(y)$ of Y_1.

b) Write down the moment generating function of Y_1.

c) Write down the mean of Y_1.

d) Let Y_2 be the time (in minutes) until the entrance of the 2^{nd} car into this stretch of freeway. Obtain $P(Y_2 > .3)$ and $P(Y_2 > .6 \mid Y_2 > .3)$.

e) Explain the reason for the equality (*) displayed in the statement of the problem.

2. Let X and Y have the joint probability density function

$$f_{X,Y}(x,y) = \begin{cases} x^2 y e^{-xy} & \text{for } 1 \le x \le 2 \text{ and } 0 < y < \infty, \\ 0 & \text{elsewhere.} \end{cases}$$

(a) Find the probability density function $f_X(x)$ of X.

(b) Find $\text{cov}(X,Y)$.

(c) Are X and Y independent?

(d) Let $U = XY$ and $V = X$. Find the joint probability density function $f_{U,V}(u,v)$ of U and V.

(e) Are U and V independent?

3. If X is a random variable whose logarithm is normally distributed (that is, $log(X) \sim N(\mu, \sigma^2)$), then X is said to have a lognormal distribution.

   (a) Let $Y \sim N(\mu, \sigma^2)$. Derive the pdf of $X = e^Y$.

   (b) Compute the expected value and variance of X using properties of the Normal distribution. (Hint: Use the mgf of Y. No integration is required.)

   (c) This distribution is popular in modeling applications, when the random variable is skewed to the right. Suppose the random variable X is income in dollars. Then $log(X)$ is a normal random variable. Fit the lognormal distribution to data from $n = 5$ employees in a company: 10,000, 25,000, 30,000, 70,000, 100,250.

      i. For these data describe how you would estimate the mean and variance of the random variable X using the lognormal distribution. Give specific formulas for estimating μ and σ^2. Compute estimates of $E[X]$, $Var(X)$, and $SD(X)$.

      ii. What is the numerical difference between using estimators $\bar{x}$ and $\hat{E}[X]$?

      iii. Find an estimate of the 95^{th} percentile of incomes from this company. Explain any properties of estimators you have used in developing your estimator. Does your numerical estimate seem reasonable?

4. The following sample of $n = 25$ observations was taken from an exponential distribution with unknown rate τ. That is, the population density function is $f(x) = \tau e^{-\tau x}$, for $x > 0$.

```
0.105, 0.145, 0.134, 0.023, 0.104, 0.074, 0.417, 0.097, 0.125, 0.346,
0.056, 0.559, 0.026, 0.494, 0.044, 0.051, 0.595, 0.098, 0.048, 0.571,
0.092, 0.064, 0.193, 0.250, 0.042.
```

The mean of these data is 0.19012, and their standard deviation is 0.18850.

a. ***Point estimates.*** (i) For general n, derive the maximum likelihood estimator (MLE) of τ. (ii) Show that multiplying the MLE by $(n-1)/n$ gives an unbiased estimator of τ. (iii) Briefly explain why this multiple of the MLE is uniformly minimum variance unbiased (UMVUE) for τ. (iv) For the data provided, compute the numerical values of the MLE and the UMVUE.

b. ***Interval estimates.*** (i) For general n, using the pivotal method and the sufficient statistic S for τ, find a 90% confidence interval for τ; give the general formula and (based on the data provided) the numerical limits. What is the pivotal quantity and what is its distribution? [Hint: Recall that if X is gamma with shape parameter κ and scale parameter 1, then $2X$ is chi-squared with 2κ degrees of freedom.] (ii) Briefly describe how to find a 90% parametric bootstrap confidence interval for τ, based on the UMVUE.

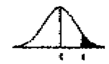
c. ***Hypothesis test.*** In terms of the sufficient statistic S, state the rejection region of a most powerful size $\alpha = 5\%$ test of the hypothesis H_0: $\tau = 5$ against H_1: $\tau = 10$. Do you reject H_0 for the data provided? How would you find the power of this test?

d. ***Some computations.*** Below is an S-Plus program (left panel) and some output from it (right). (i) Some of the code is directly relevant to earlier parts of this question. Specifically identify these lines of code and explain briefly. (ii) Some of the code explores alternate procedures. Explain.

```
m <- 10000;  n <- 25;  tau <- 5
x <- rexp(m*n, rate=tau)
DTA <- matrix(x, nrow=m)

x.mean <- rowMeans(DTA)
x.sd <- sqrt(rowVars(DTA))
t1 <- 1/x.mean
t2 <- (n-1)/(n*x.mean); mean(t2)
t3 <- (n-2)/(n*x.mean)
t4 <- (n-3)/(n*x.sd)

round(c(mean(t1), mean((t1-5)^2)), 3)
round(c(mean(t2), mean((t2-5)^2)), 3)
round(c(mean(t3), mean((t3-5)^2)), 3)
round(c(mean(t4), mean((t4-5)^2)), 3)

s <- n*0.19012
adj <- (1/tau)*(n-1)/s
round(adj*quantile(t2, c(.05, .95)), 2)
round(qgamma(c(.05,.95), n, 1)/s, 2)
```

```
> round(c(mean(t1), mean((t1 - 5)^2)), 3)
[1] 5.204 1.222
> round(c(mean(t2), mean((t2 - 5)^2)), 3)
[1] 4.996 1.088
> round(c(mean(t3), mean((t3 - 5)^2)), 3)
[1] 4.788 1.045
> round(c(mean(t4), mean((t4 - 5)^2)), 3)
[1] 4.885 1.820

> s <- n * 0.19012
> adj <- ((1/tau) * (n - 1))/s
> round(adj * quantile(t2, c(0.05, 0.95)), 2)
   5%  95%
 3.61 6.95

> round(qgamma(c(0.05, 0.95), n, 1)/s, 2)
[1] 3.66 7.10
```

Upper-tail Areas for the Normal Curve



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.00	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.10	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.20	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.30	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.40	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.50	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.60	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.70	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.80	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.90	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.00	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.10	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.20	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.30	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.40	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.50	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.60	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.70	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.80	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.90	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.00	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.10	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.20	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.30	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.40	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.50	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.60	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.70	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.80	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.90	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.00	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

z	Area
3.500	.00023263
4.000	.00003167
4.500	.00000340
5.000	.00000029

Percentage Points of the t Distribution

df	$a = .1$	$a = .05$	$a = .025$	$a = .01$	$a = .005$	$a = .001$
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
60	1.296	1.671	2.000	2.390	2.660	3.232
120	1.289	1.658	1.980	2.358	2.617	3.160
240	1.285	1.651	1.970	2.342	2.596	3.125
inf.	1.282	1.645	1.960	2.326	2.576	3.090

100 × γth Percentiles $\chi^2_\gamma(v)$ of the chi-square distribution with v degrees of freedom

$$\gamma = \int_0^{\chi^2_\gamma(v)} h(y;\, v)\, dy$$

v	γ = 0.005	0.010	0.025	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.975	0.990	0.995	0.999
1					0.02	0.10	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75	20.52
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.09	21.96	26.12
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.92	24.72	26.76	31.26
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.49	24.77	27.59	30.19	33.41	35.73	40.79
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.72	27.20	30.14	32.85	36.19	35.58	43.82
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.83	28.41	31.41	34.17	37.57	40.00	45.32
21	8.03	8.90	10.28	11.59	13.24	16.34	20.34	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	8.64	9.54	10.98	12.34	14.04	17.24	21.34	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	9.26	10.20	11.69	13.09	14.85	18.14	22.34	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	9.89	10.86	12.40	13.85	15.66	19.04	23.34	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	10.52	11.52	13.12	14.61	16.47	19.94	24.34	29.34	34.38	37.65	40.65	44.31	46.93	52.62
30	13.79	14.95	16.79	18.49	20.60	24.48	29.34	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	20.71	22.16	24.43	26.51	29.05	33.66	39.34	45.62	51.80	55.76	59.34	63.69	66.77	73.40
50	27.99	29.71	32.36	34.76	37.69	42.94	49.33	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	35.53	37.48	40.48	43.19	46.46	52.29	59.33	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	43.28	45.44	48.76	51.74	55.33	61.70	69.33	77.58	85.53	90.53	95.02	100.42	104.22	112.32
80	51.17	53.54	57.15	60.39	64.28	71.14	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
90	59.20	61.75	65.65	69.13	73.29	80.62	89.33	98.64	107.56	113.14	118.14	124.12	128.30	137.21
100	67.33	70.06	74.22	77.93	82.36	90.13	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45