STUDENT NAME: STUDENT ID#

McGILL UNIVERSITY
FACULTY OF SCIENCE

FINAL EXAMINATION

MATH324

STATISTICS

Examiner: Professor M. Asgharian
Associate Examiner: Professor D. Wolfson

Date: Tuesday, December 19, 2006
Time: 2:00 P.M. - 5:00 P.M.

INSTRUCTIONS

Answer ONLY 8 questions.
This is a closed book exam.
Calculators are permitted.
Answer directly on the exam.
Dictionaries are allowed
Tables have been provided.
One page of formulas is permitted.

Questions	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	

This exam comprises the cover page, 9 pages of questions, 1 blank page for rough work and 5 pages of tables.

1. The reading on a voltage meter connected to a test circuit is uniformly distributed over the interval $(\theta, \theta + 1)$, where θ is the true but unknown voltage of the circuit. Suppose that $Y_1, \cdots, Y_n$ denotes a random sample of such readings.

   (a) Show that $\overline{Y}$ is a biased estimator of θ, and compute the bias. (*3 marks*)

   (b) Find a function of $\overline{Y}$ that is an unbiased estimator of θ. (*3 marks*)

   (c) Find $MSE(\overline{Y})$ when $\overline{Y}$ is used as an estimator of θ. (*4 marks*)

2. Suppose $X_1, X_2, \cdots, X_n$ is a random sample of size n from a Poisson distribution with mean λ.

   (a) Using Tchebysheff's inequality show that $\overline{X}$ is a consistent estimator for λ. (*3 marks*)

   (b) Define $W_n = \frac{\overline{X}-\lambda}{\sqrt{\overline{X}/n}}$. Show that the distribution of W_n converges to a standard normal distribution. (*4 marks*)

   (c) Use W_n and the result in part (b) to derive the formula for an approximate 95% confidence interval for λ. (*3 marks*)

3. Suppose that $X_1, \cdots, X_m$, representing yields per acre for corn variety A, constitute a random sample from a normal distribution with mean μ_1 and variance σ^2. Also, $Y_1, \cdots, Y_n$, representing yields for corn variety B, constitute a random sample from a normal distribution with mean μ_2 and variance σ^2. Suppose, further, that the Xs and Ys are independent.

   (a) Find a sufficient statistic for (μ_1, μ_2, σ^2). *(5 marks)*

   (b) Find the maximum likelihood estimator for σ^2. *(5 marks)*

4. A two-stage clinical trial is planned for testing $H_0 : p = 0.15$ versus $H_A : p > 0.15$, where p is the proportion of responders among patients who were treated by the protocol treatment. At the first stage, 20 patients are accrued and treated. If 5 or more responders are observed among the (first) 20 patients, H_0 is rejected, the study is terminated, and no more patients are accrued. Otherwise, another 20 patients will be accrued and treated in the second stage. If a total of 8 or more responders are observed among the 40 patients accrued in the two stages (20 in the first stage and 20 more in the second stage), then H_0 is rejected.

   (a) Use the binomial table to find the numerical value of α for this testing procedure. (*5 marks*)

   (b) Use the binomial table to find the probability of rejecting the null hypothesis when using this rejection region if $p = 0.40$. (*2 marks*)

   (c) For the rejection region defined above, find β if $p = 0.40$. (*3 marks*)

5. Let $Y_1, Y_2, \cdots, Y_n$ denote a random sample from a Bernoulli-distributed population with parameter p. That is,

$$P(Y = y; p) = p^y(1-p)^{1-y}, \qquad y = 0, 1.$$

(a) Suppose that we are interested in testing $H_0 : p = p_0$ versus $H_a : p = p_a$, for any $p_0 < p_a$. Show that

$$\frac{L(p_0)}{L(p_a)} = \left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right)^{\sum_{i=1}^n y_i} \left(\frac{1-p_0}{1-p_a}\right)^n.$$

Argue that $L(p_0)/L(p_a) < k$ if and only if $\sum_{i=1}^n y_i > k^*$ for some constant k^*. Give the rejection region for the most powerful test of size α, H_0 versus H_a. *(6 marks)*

(b) Recall that $\sum_{i=1}^n Y_i$ has a binomial distribution with parameters n and p. Indicate how to determine the values of any constants contained in the rejection region derived in part (a). *(4 marks)*

6. Suppose that an engineer wishes to compare the number of complaints per week filed by union stewards for two different shifts at a manufacturing plant. One hundred independent observations on the number of complaints gave means $\overline{x} = 20$ for shift 1 and $\overline{y} = 22$ for shift 2. Assume that the number of complaints per week on the ith shift has a Poisson distribution with mean θ_i, for $i = 1, 2$. Use the likelihood ratio method to test $H_0 = \theta_1 = \theta_2$. Find an upper bound for the p-value. (*Hint*: We know that $-2\ln(\lambda) \overset{asmp}{\sim} \chi^2_{\dim(\Omega)-\dim(\Omega_0)}$, where Ω and Ω_0 are, respectively, the whole parameter space and the parameter space under H_0.) You may assume that the number of complaints on the two different shifts are independent. (*10 marks*)

7. The birth weight, x, and the increase in weight between the 70th and 100th day of life, y, expressed as a percentage of the birth weight has been measured on 32 babies. The data suggests an association between the two variables in a negative direction. This seems quite plausible: when the birth weight is low the subsequent rate of growth, *relative to the birth weight*, would be expected to be high, and vice versa. The trend seems reasonably linear. The regression line fitted using the least square method is $\hat{y} = 167.87 - 0.8643x$ with $\overline{x} = 111.75$, $V(b) = (0.1757)^2$, $S_{XX} = 10262$ and $SS_{Total} = 17168.47$.

   (a) Test the hypothesis that the slope is different from zero at the significance level of 0.05. (*3 marks*)

   (b) Find a 95% confidence interval for the mean percentage increase in birth weight at $x = 110$. (*3 marks*)

   (c) Find the coefficient of determination, R^2, and comment on the quality of the fit. (*4 marks*)

8. It is believed that women in the postmenopausal phase of life suffer from calcium deficiency. This phenomenon is associated with the relatively high proportion of bone fractures for women in that age group. Is this calcium deficiency associated with an estrogen deficiency, a condition that occurs after menopause? To investigate this theory, Richardson, Wahner, Melton and Riggs (1984, *New England Journal of Medicine 311(20),pp. 1273-75*) compared the bone mineral density in three groups of women.

   The first group of 14 women had undergone oophorectomy during young adult womanhood and had lived for a period of 15 to 25 years with an estrogen deficiency. A second group, identified as premenopausal, were approximately the same age (approximately 50 years) as the oophorectomy group except that the women had never suffered a period of estrogen deficiency. The third group of 14 women were postmenopausal and had suffered an estrogen deficiency for an average of 20 years. The mean and standard deviation of the mean for the three samples of lumbar spine bone density measurements-14 measurements in each sample, one for each subject-are recorded in the following table.

	Groups		
Statistic	Oophorectomized	Premenopausal	Postmenopausal
Mean	.93	1.21	.92
Standard Error	.04	.03	.04

   Is there sufficient evidence to permit us to conclude that the mean bone-density measurements differ for the three groups of women? What is the p-value associated with your test? What would you conclude at the α=.05 level? (*10 marks*)

9. Brush et al.(1985, *New England Journal of Medicine*) conducted a study suggesting that the initial electrocardiogram (ECG) of a suspected heart attack victim can be used to predict in-hospital complications of an acute nature. The study included 469 patients with suspected myocardial infarction (heart attack). Each patient was categorized according to whether their initial ECG was positive or negative and whether the person suffered life-threatening complications subsequently in the hospital. The results are summarized in the following table.

	Subsequent In-Hospital Life-Threatening Complications		
ECG	No	Yes	Total
Negative	166	1	167
Positive	260	42	302
Total	420	43	469

(a) Is there sufficient evidence to indicate that whether or not a heart attack patient suffers complications depends on the outcome of the initial ECG? Test using $\alpha = .05$. (*7 marks*)

(b) Give bounds for the observed significance level. (*3 marks*)

Table 1. (*Continued*)

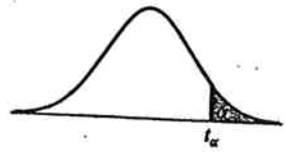
(c) $n = 15$

a	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99	*a*
0	.860	.463	.206	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.990	.829	.549	.167	.035	.005	.000	.000	.000	.000	.000	.000	.000	1
2	1.000	.964	.816	.398	.127	.027	.004	.000	.000	.000	.000	.000	.000	2
3	1.000	.995	.944	.648	.297	.091	.018	.002	.000	.000	.000	.000	.000	3
4	1.000	.999	.987	.836	.515	.217	.059	.009	.001	.000	.000	.000	.000	4
5	1.000	1.000	.998	.939	.722	.403	.151	.034	.004	.000	.000	.000	.000	5
6	1.000	1.000	1.000	.982	.869	.610	.304	.095	.015	.001	.000	.000	.000	6
7	1.000	1.000	1.000	.996	.950	.787	.500	.213	.050	.004	.000	.000	.000	7
8	1.000	1.000	1.000	.999	.985	.905	.696	.390	.131	.018	.000	.000	.000	8
9	1.000	1.000	1.000	1.000	.996	.966	.849	.597	.278	.061	.002	.000	.000	9
10	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.164	.013	.001	.000	10
11	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.352	.056	.005	.000	11
12	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.602	.184	.036	.000	12
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.833	.451	.171	.010	13
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.794	.537	.140	14

(d) $n = 20$

a	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99	*a*
0	.818	.358	.122	.012	.001	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.983	.736	.392	.069	.008	.001	.000	.000	.000	.000	.000	.000	.000	1
2	.999	.925	.677	.206	.035	.004	.000	.000	.000	.000	.000	.000	.000	2
3	1.000	.984	.867	.411	.107	.016	.001	.000	.000	.000	.000	.000	.000	3
4	1.000	.997	.957	.630	.238	.051	.006	.000	.000	.000	.000	.000	.000	4
5	1.000	1.000	.989	.804	.416	.126	.021	.002	.000	.000	.000	.000	.000	5
6	1.000	1.000	.998	.913	.608	.250	.058	.006	.000	.000	.000	.000	.000	6
7	1.000	1.000	1.000	.968	.772	.416	.132	.021	.001	.000	.000	.000	.000	7
8	1.000	1.000	1.000	.990	.887	.596	.252	.057	.005	.000	.000	.000	.000	8
9	1.000	1.000	1.000	.997	.952	.755	.412	.128	.017	.001	.000	.000	.000	9
10	1.000	1.000	1.000	.999	.983	.872	.588	.245	.048	.003	.000	.000	.000	10
11	1.000	1.000	1.000	1.000	.995	.943	.748	.404	.113	.010	.000	.000	.000	11
12	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.032	.000	.000	.000	12
13	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.087	.002	.000	.000	13
14	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.196	.011	.000	.000	14
15	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.370	.043	.003	.000	15
16	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.589	.133	.016	.000	16
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.794	.323	.075	.001	17
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.931	.608	.264	.017	18
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.988	.878	.642	.182	19

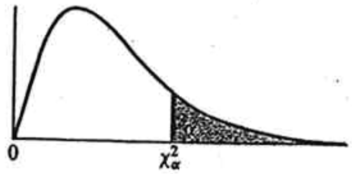
Table 5. Percentage points of the *t* distributions



$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	d.f.
3.078	6.314	12.706	31.821	63.657	1
1.886	2.920	4.303	6.965	9.925	2
1.638	2.353	3.182	4.541	5.841	3
1.533	2.132	2.776	3.747	4.604	4
1.476	2.015	2.571	3.365	4.032	5
1.440	1.943	2.447	3.143	3.707	6
1.415	1.895	2.365	2.998	3.499	7
1.397	1.860	2.306	2.896	3.355	8
1.383	1.833	2.262	2.821	3.250	9
1.372	1.812	2.228	2.764	3.169	10
1.363	1.796	2.201	2.718	3.106	11
1.356	1.782	2.179	2.681	3.055	12
1.350	1.771	2.160	2.650	3.012	13
1.345	1.761	2.145	2.624	2.977	14
1.341	1.753	2.131	2.602	2.947	15
1.337	1.746	2.120	2.583	2.921	16
1.333	1.740	2.110	2.567	2.898	17
1.330	1.734	2.101	2.552	2.878	18
1.328	1.729	2.093	2.539	2.861	19
1.325	1.725	2.086	2.528	2.845	20
1.323	1.721	2.080	2.518	2.831	21
1.321	1.717	2.074	2.508	2.819	22
1.319	1.714	2.069	2.500	2.807	23
1.318	1.711	2.064	2.492	2.797	24
1.316	1.708	2.060	2.485	2.787	25
1.315	1.706	2.056	2.479	2.779	26
1.314	1.703	2.052	2.473	2.771	27
1.313	1.701	2.048	2.467	2.763	28
1.311	1.699	2.045	2.462	2.756	29
1.282	1.645	1.960	2.326	2.576	inf.

From "Table of Percentage Points of the *t*-Distribution." Computed by Maxine Merrington, *Biometrika*, Vol. 32 (1941), p. 300. Reproduced by permission of Professor E. S. Pearson.

Table 6. Percentage points of the χ^2 distributions



d.f.	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$
1	0.0000393	0.0001571	0.0009821	0.0039321	0.0157908
2	0.0100251	0.0201007	0.0506356	0.102587	0.210720
3	0.0717212	0.114832	0.215795	0.351846	0.584375
4	0.206990	0.297110	0.484419	0.710721	1.063623
5	0.411740	0.554300	0.831211	1.145476	1.61031
6	0.675727	0.872085	1.237347	1.63539	2.20413
7	0.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

Table 7. (*Continued*)

F_α

Denominator d.f.	α	1	2	3	4	5	6	7	8	9
		Numerator d.f.								
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

Table 7. (*Continued*)

F_α

Denominator d.f.	Numerator d.f.									
	α	1	2	3	4	5	6	7	8	9
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.005	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.005	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.005	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52