STUDENT NAME: STUDENT ID#

McGILL UNIVERSITY
FACULTY OF SCIENCE

FINAL EXAMINATION

MATH323B

PROBABILITY THEORY

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

Date: Monday, April 18, 2005
Time: 2:00 P.M. - 5:00 P.M.

INSTRUCTIONS

Answer ONLY 8 questions.
Calculators are permitted.
Answer directly on exam
Dictionaries are allowed

Questions	Marks
1	
2	
3	
4	
5	
5	
6	
7	
8	
9	
10	

This exam comprises the cover, 10 pages of questions, 2 pages of formulas and 1 page of tables.

1. Suppose that two balanced dice are tossed repeatedly and the sum of the two uppermost faces is determined on each toss. What is the probability that we obtain a sum of 4 before we obtain a sum of 7?

2. Suppose 5% of all people filing the long income tax form seek deductions that they know are illegal, while 2% incorrectly list deductions because they are unfamiliar with income tax regulations. Of the 5% who are guilty of cheating, 80% will deny knowledge of the error if confronted by an investigator. If the filer of the long form is confronted with an unwarranted deduction and he or she denies the knowledge of the error, what is the probability that he or she is guilty?

3. It is known that 5% of the members of a population have disease A, which can be discovered by a blood test. Suppose that N (a large number) people are to be tested. This can be done in two ways: (1) Each person is tested separately; or (2) the blood samples of k people are pooled together and analyzed. (Assume that $N = nk$, with n an integer.) If the test is negative, all of them are healthy (that is, just this one test is needed). If the test is positive, each of the k persons must be tested separately (that is, a total of $k + 1$ are needed).

   (a) For fixed k what is the expected number of tests needed in (2)?

   (b) Find the k that will minimize the expected number of tests in (2)?

   (c) If k is selected as in (b), on average how many tests does (2) save in comparison with (1)?

4. Suppose that plants of a particular species are randomly dispersed over an area, so that the number of plants in a given area follows a Poisson distribution with a mean density of λ plants per unit area. If a plant is randomly selected in this area, find the probability density function of the distance to the *nearest* neighboring plant. [*Hint:* If R denotes the distance to the nearest neighbor, then $P(R > r)$ is the same as the probability of seeing no plants in a circle of radius r.]

5. Let Y denote a random variable with probability density function given by

$$f(y) = (1/2)e^{|y|}, \qquad -\infty < y < \infty \quad .$$

Find the moment-generating function of Y, and use it to find $E(Y)$.

6. The type of medical care a patient receives may vary with the age of the patient. Here is the joint probability mass function associated with data obtained in a large study of women who had a breast lump investigated, whether or not each woman received a mammogram and a biopsy when the lump was discovered. Define

$$Y_1 = \begin{cases} 0, & \text{if the patient is under 65 years of age } (< 65), \\ 1, & \text{if the patient's age is above 65 years } (\geq 65). \end{cases}$$

and

$$Y_2 = \begin{cases} 0, & \text{if the test was not done,} \\ 1, & \text{if the test was done.} \end{cases}$$

Then the following table provides $P_{ij} = P(Y_1 = i,\ Y_2 = j)$:

		Y_1	
		0	1
Y_2	0	0.124	0.190
	1	0.321	0.365

.

(a) Find the marginal probability functions for Y_1 and Y_2.

(b) Find the conditional probability function for Y_2 given $Y_1 = 1$. Are Y_1 and Y_2 are independent? Why?

(c) Were the tests omitted on older patients more or less frequently than would be the case if testing were independent of age?

7. The National Fire Incident Reporting Service stated that, among residential fires, 73% are in family homes, 20% are in apartments, and 7% are in other types of dwellings. The typical cost of damages caused by a fire in a family home is \$20,000. Comparable costs for an apartment fire and for fire in other dwelling types are \$10,000 and \$2,000, respectiely. Suppose four residential fires are independently reported on a single day,

   (a) Find the expected value and variance of the total damage cost.

   (b) What is the probability that two are in family homes, one is in an apartment, and one is in another type of dwelling?

8. The weight (in pounds) of "medium-size" watermelons is normally distributed with mean 15 and variance 4. A packing container for several melons has a nominal capacity of 140 pounds. What is the maximum number of melons that should be placed in a single packing container if the nominal weight limit is to be exceeded only 5% of the time? Give reasons for your answer.

9. Let $X_1, X_2, \cdots, X_n$ be independent χ^2-distributed random variables, each with 1 degree of freedom. Define $Y = \sum_{i=1}^{n} X_i$.

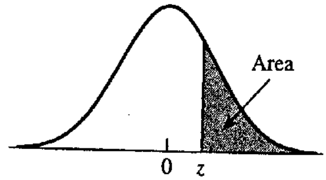
(a) Use the method of moment generating functions to show that Y has a χ^2 distribution with n degrees of freedom.

(b) Use the preceding representation of Y as the sum of the $X's$ to show that $Z = (Y - n)/\sqrt{2n}$ has an asymptotic standard normal distribution.

(c) A machine in a heavy-equipment factory produces steel rods of length Y, where Y is a normally distributed random variable with mean $\mu = 6$ inches and variance $\sigma^2 = 0.2$ inches^2. The cost C of repairing a rod that is not exactly 6 inches in length is proportional to the square of the error and is given, in dollars, by $C = 4(Y - \mu)^2$. If 50 rods with independent lengths are produced in a given day, approximate the probability that the total cost for repairs for that day exceeds \$48.

10. An urn contains a total of N balls, some black and some white. Samples are drawn from the urn, m balls at a time ($m \leq N$). After drawing each sample, the black balls are returned to the urn, while the white balls are replaced by black balls and then returned to the urn. If the number of white balls in the urn is i, we say that the "system" is in the state ε_i. Show that the random process described by this model is a Markov chain (imagine that samples are drawn at the times $t = 1, 2, \cdots$, and that the system has some initial distribution). Find the corresponding transition probabilities p_{ij} $(i, j = 0, 1, \cdots, N)$.

Table 4. Normal curve areas
Standard normal probability in right-hand tail
(for negative values of z areas are found by symmetry)



	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000 233									
4.0	.000 031 7									
4.5	.000 003 40									
5.0	.000 000 287									

From R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

Continuous distributions

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Uniform	$f(y) = \frac{1}{\theta_2 - \theta_1}; \theta_1 \le y \le \theta_2$	$\frac{\theta_1 + \theta_2}{2}$	$\frac{(\theta_2 - \theta_1)^2}{12}$	$\frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$
Normal	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$	μ	σ^2	exp $\mu t + \frac{t^2\sigma^2}{2}\Big)$
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta}; \quad \beta > 0$ $0 < y < \infty$	β	β^2	$(1 - \beta t)^{-1}$
Gamma	$f(y) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1} e^{-y/\beta};$ $0 < y < \infty$	$\alpha\beta$	$\alpha\beta^2$	$(1 - \beta t)^{-\alpha}$
Chi-square	$f(y) = \frac{(y)^{(v/2)-1} e^{-y/2}}{2^{v/2}\Gamma(v/2)};$ $y^2 > 0$	v	$2v$	$(1 - 2t)^{-v/2}$
Beta	$f(y) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1}(1-y)^{\beta-1};$ $0 < y < 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	does not exist in closed form

Discrete distributions

Distribution	Probability Function	Mean	Variance	Moment-Generating Function
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y}$; $y = 0, 1, \ldots, n$	np	$np(1-p)$	$[pe^t + (1-p)]^n$
Geometric	$p(y) = p(1-p)^{y-1}$; $y = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hypergeometric	$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$; $y = 0, 1, \ldots, n$ if $n \le r$, $y = 0, 1, \ldots, r$ if $n > r$	$\frac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$	
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!}$; $y = 0, 1, 2, \ldots$	λ	λ	$\exp[\lambda(e^t - 1)]$
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$; $y = r, r+1, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$