NATIONAL UNIVERSITY OF SINGAPORE

EXAMINATION

ST2334 PROBABILITY AND STATISTICS

(Semester 1: AY 2008–2009)

November 2008 — Time Allowed : **2 Hours**

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **FIVE (5)** questions and comprises **ELEVEN (11)** printed pages.
2. Answer **ALL** the questions correctly for TOTAL 60 marks.
3. Read the questions **CAREFULLY**, and label some important quantities clearly at your own convenience.
4. This is an **OPEN BOOK** examination.
5. **Decimal answers with more than 4 decimal places** should be given to at least **FOUR (4)** decimal places.

Matriculation No: _____

Question	1	2	3	4	5	Total
Max. marks	7	14	6	15	18	60
Marks scored						

1. **[7 marks]** Suppose that past data report that approximately 60% of students in the module ST2334 attend lectures. In a random group of 15 students who officially enrol in the module this semester, let x denote the number of students being present in a lecture.

   (a) Within the group of 15 students, what is the expected number of absentees in the last lecture in this semester?

   (b) What is the largest value of c for which $P(x \leq c) \leq 0.01$?

   (c) Suppose that there are 50 students who officially enrol in ST2334 this semester. To check attendance in each lecture, suppose that 15 names of the students are randomly selected and read out, how likely is it that there are strictly fewer than 4 students being present in a lecture?

2. **[14 marks]** Accident records collected by an automobile insurance company give the following information: The probability that an insured driver has no automobile accident is 0.85; if an accident has occurred, the damage to the vehicle amounts to a total loss with probability 0.08, or 60% of its market value with probability 0.12, or 20% of its market value in other situations. Suppose that the company only insures \$22,000 priced cars. Define

C : the premium (in \$) which the company charges

x : the company's gain (in \$) by charging \$$C$ as premium

(a) Give the four possible values of the random variable x in terms of C.

(b) Justify that there is 1.2% of the time the company's gain is given by $C - 22,000$.

(c) Fill in the following table which shows the probability distribution of x.

x	$C - 22,000$			
$p(x)$	0.012			

(d) Show that the company should charge on a \$22,000 priced car $C = \$1,200$ as the premium so that the expected gain by the company is \$171.4.

170.4

(e) By charging the premium $C = \$1,200$, what is the standard deviation of the company's gain?

3. [**6 marks**] A wildlife service wishes to estimate, μ, the mean number of days of hunting per hunter for all hunters licensed in a region during a given season. 50 hunters are interviewed. An average number of 4.5 days of hunting per hunter is recorded, and the sample standard deviation equals 10 days.

   (a) Construct an approximately 99% confidence interval for μ.

   (b) In order to achieve a bound on the error of estimation as 2 hunting days with 99% degree of confidence, at least how many more hunters must be included in the study?

4. **[15 marks]** The maximum load (with a generous safety factor) for the elevator in an office building is 2,000 pounds. The relative frequency distribution of the weights of all men and women using the elevator is mound-shaped, with mean $\mu = 150$ pounds and standard deviation $\sigma = 35$ pounds. Assume that there are n passengers in the elevator. Give your answer in terms of n when it is necessary.

   (a) What is the expected total weight of passengers in the elevator? What is the standard deviation?

   (b) Describe the probability distribution of the total weight of n passengers in the elevator. Explain.

(c) Show that there is less than 1% of the time that the total weight of 11 passengers in the elevator exceeds the maximum load.

(d) Show that the probability that the total weight of 12 passengers in the elevator exceeds the maximum load is 0.0495.

(e) What is the limit of the number of people on the elevator you would recommend if you want to keep the chance that the total weight exceeds the maximum load to be less than 1%?

5. **[18 marks]** The proportion p of Singapore households in Housing & Development Board (HDB) flats with two or more television (TV) sets is of interest. Among 100 randomly selected households, 35 of them have two or more TV sets.

   (a) In usual practice, one would use the sample proportion of households with two or more TV sets as an unbiased estimator of p. Explain in this context how good it is by using $\hat{p}$ as an unbiased estimator.

   (b) Give an approximately 90% confidence interval for p. Do you need to check any assumption?

   (c) Explain why you are about 90% sure that p lies within the interval in part (b).

(d) A survey conducted 5 years ago indicates that 30% of Singapore households in HDB flats have two or more television (TV) sets. Someone argues that this may not be valid anymore, and decides to carry out a hypothesis test.

(i) State the null and the alternative hypotheses.

(ii) Give the observed value of the test statistic **and** the p-value for the data in the previous page.

(iii) Suppose that the level of significance is 5%. State the conclusion of the test.

Cumulative Binomial Probability Table:

(a) $n = 5$

	p												
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	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000	0
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000	1
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000	2
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001	3
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049	4

(b) $n = 10$

	p												
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	.904	.599	.349	.107	.028	.006	.001	.000	.000	.000	.000	.000	.000	0
1	.996	.914	.736	.376	.149	.046	.011	.002	.000	.000	.000	.000	.000	1
2	1.000	.988	.930	.678	.383	.167	.055	.012	.002	.000	.000	.000	.000	2
3	1.000	.999	.987	.879	.650	.382	.172	.055	.011	.001	.000	.000	.000	3
4	1.000	1.000	.998	.967	.850	.633	.377	.166	.047	.006	.000	.000	.000	4
5	1.000	1.000	1.000	.994	.953	.834	.623	.367	.150	.033	.002	.000	.000	5
6	1.000	1.000	1.000	.999	.989	.945	.828	.618	.350	.121	.013	.001	.000	6
7	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070	.012	.000	7
8	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264	.086	.004	8
9	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651	.401	.096	9

(c) $n = 15$

	p												
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	.995	.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

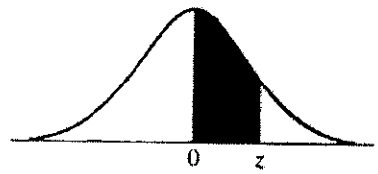

Table of normal curve areas

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

SOURCE: This table is abridged from Table 1 of *Statistical Tables and Formulas* by A. Hald (New York: Wiley, 1952). Reproduced by permission of A. Hald and the publisher, John Wiley & Sons, Inc.

[END OF PAPER]