NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS & STATISTICS

SEMESTER 2 2006–2007

ACE2013

Statistics for Marketing and Management

Specimen Paper

Time allowed: 2 hours

Credit will be given for all answers to questions in Section A, and for the best TWO answers to questions in Section B. No credit will be given for other answers and students are strongly advised not to spend time producing answers for which they will receive no credit.

Marks for each question are indicated. However you are advised that marks indicate the relative weight of individual questions, they do not correspond directly to marks on the University scale.

There are SIX questions in Section A and FOUR questions in Section B.

Graph paper will be provided. Statistical tables and formulae are provided at the end of this exam paper.

SECTION A

A1. The data below represent the weekly sales at a Mercedes–Benz car franchise.

$$2 \quad 5 \quad 4 \quad 3 \quad 26 \quad 4 \quad 1 \quad 2 \quad 5 \quad 3$$

(a) Calculate the mean and standard deviation for these data.

(b) Find the median.

(c) Which measure of average – the mean or the median – would you prefer for these data? Explain.

(d) Suggest an alternative measure of *spread*.

[**7 marks**]

A2. Sixty percent of calls to the HSBC Complaints line in Newcastle are immediately put "on hold", where the caller is subjected to crap classical music. Between 9am and 10am this morning, twenty people called to make complaints.

(a) What probability distribution might be reasonable to use to model the number of calls put "on hold"?

(b) What is the expected number and standard deviation of the number of calls put on hold?

(c) What is the probability that, between 9am and 10am, at least one person is put "on hold"?

[**7 marks**]

A3. *Choctastic!* produce large chocolate chip cookies. When their production process is working satisfactorily the weights of individual cookies have a weight of 98g with a standard deviation of 4g.

(a) Customers are entitled to a refund if a cookie weighs less than 93g. What percentage of customers can we expect to be entitled to a refund?

(b) What percentage of cookies will weigh between 93g and the stated weight of 98g?

[**7 marks**]

A4. *Choctastic!* also produce bars of Turkish Delight. When their "gooing" machine is working properly, these bars have an average weight of 30g. For quality control, a sample of 20 bars is taken during the course of a day, giving an average weight of 29.7g with a standard deviation of 1.2g.

(a) Construct a 90% confidence interval for the average weight of a bar of Turkish Delight.

(b) Would you say the "gooing" machine is working properly at the minute? Justify your answer.

(c) Would a 95% confidence interval be narrower or wider than that obtained in part (a)? Why? *[Do not perform any further calculations]*

[**8 marks**]

A5. The latest mobile phone on the market, the "Samsung Brick", is now sold in most European countries. A random sample of phones bought in the U.K. and mainland Europe gave the following results:

U.K.	$n_1 = 17$	$\bar{x}_1 = £175.20$	$s_1 = £17.35$
Mainland Europe	$n_2 = 14$	$\bar{x}_2 = £160.59$	$s_2 = £25.67$

Perform an appropriate hypothesis test to determine whether there is a difference between the mean retail price of the Samsung Brick in the U.K. and mainland Europe. *[Hint: The pooled standard deviation is £21.48]*

[**9 marks**]

A6. The Personnel Manager of a company believes that monthly paid staff take more time off work through sickness than those staff who are paid weekly (and do not belong to the company sickness scheme). To test this theory, the sickness records for 531 randomly selected employees who have been in continuous employment for the past year were analysed. These have been summarised in the table below.

	Number of days off sick			
	Less than 5	5 to 10	More than 10	**Total**
Monthly paid	85	41	11	137
Weekly paid	136	152	106	394
Total	221	193	117	531

Is there evidence of an association between type of employee and number of days off sick?

[**12 marks**]

SECTION B

B7. Shown below is a grouped frequency table showing the amount spent (X, in pounds) at Morrison's in Jarrow by male and female customers.

Amount spent (X £)	Male	Female
$0 \leq x < 10$	21	4
$10 \leq x < 20$	33	12
$20 \leq x < 30$	27	18
$30 \leq x < 40$	15	25
$40 \leq x < 50$	10	47
$50 \leq x < 60$	6	22
$60 \leq x < 70$	2	19
$70 \leq x < 80$	0	11
$80 \leq x < 90$	1	5
$90 \leq x < 100$	0	1

(a) Use the frequency table above to estimate the mean expenditure for males and females at this branch of Morrison's. Why are your answers *approximations* of the sample means?

(b) Construct *relative percentage* frequency polygons for male and female expenditure and overlay these polygons on the same graph. Why might it be a good idea to use *relative percentage* polygons here?

(c) The standard deviations for both male and female expenditures are £6.95 and £8.05 respectively.

Assuming a Normal distribution for both male and female expenditure, find the probability that a randomly selected male and a randomly selected female will spend less than £30.

(d) Do you trust your calculations in part (c)? Comment.

[25 marks]

B8. A company has developed a new manufacturing process which they believe will revolutionise their industry. They are, however, uncertain how they should go about exploiting this advance.

Initial indications of the likely success of marketing the process are 55%, 30% 15% for "high success", "medium success" and "probable failure" (respectively). The company has three options; they can go ahead and develop the technology themselves, license it or sell the rights to it. The financial outcomes (in £ millions) for each option are given in the table below.

	"high success"	"medium success"	"failure"
Develop	80	40	–100
Licence	40	30	0
Sell	25	25	25

(a) Draw a decision tree to represent the company's problem.

(b) Calculate the *Expected Monetary Value* for all possible decisions the company may take and hence determine the optimal decision for the company.

[25 marks]

B9. The Elves Toy Company makes toy trains and dolls' prams, which use the same wheels and logo stickers.

Each train requires 8 wheels and 2 logo stickers. Each pram requires 8 wheels and 3 logo stickers. The company has 7200 wheels and 2200 logo stickers available.

To meet current demand, the company is to make at least 300 of each type of toy. The company makes and sells x trains and y prams for £20 and £25 respectively.

(a) Formulate the company's situation as a linear programming problem.

(b) Draw a suitable diagram to enable the problem to be solved graphically, indicating the feasible region and the direction of the objective line.

(c) Use your diagram to find the company's minimum and maximum total income, $£T$.

(d) Verify the solution obtained graphically in part (c) by solving this linear programming problem algebraically.

[**25 marks**]

B10. Barbara is considering buying the Dog and Duck Pub and takes a look at the sales records over the past three years. She has been given four–monthly sales figures by the current owner (Y, in thousands of pounds), and these are shown below.

	Jan–Apr	May–Aug	Sep–Dec
2004	18	28	22
2005	20	30	26
2006	22	34	28

(a) Produce a time series plot for these data, and comment.

(b) Calculate the moving averages for these data, and overlay these on your plot in part (a).

(c) Let T represent time, and let time–points $1, 2, \ldots,$ correspond to Jan–Apr (2004), May–Aug (2004), ... (respectively). Then we have the following summaries:

$$\sum t = 35 \qquad \sum y = 176.67$$

$$\sum t^2 = 203 \qquad \sum ty = 909.33$$

Use the summaries above to estimate the linear trend

$$Y = \alpha + \beta T + \epsilon.$$

(d) The (unadjusted) seasonal effects are given in the table below.

"Season"	Seasonal effects
Jan–Apr	–5.428
May–Aug	+4.571
Sep–Dec	–1.429

Obtain the *adjusted* seasonal effects.

(e) Use the linear trend equation in part (c), and the adjusted seasonal effects in part (d), to forecast sales for the current four–month period (i.e. May–Aug 2007).

(f) Why should Barbara be cautious about using this model to predict sales for Sep–Dec 2009?

[25 marks]

Statistical formulae and tables

1 Sample statistics

Lower quartile: $Q_1 = \frac{(n+1)}{4}$th smallest observation.

Median: $Q_2 = \frac{(n+1)}{2}$th smallest observation.

Upper quartile: $Q_3 = \frac{3(n+1)}{4}$th smallest observation.

Sample mean: $$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Sample variance: $$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right\}$$

2 Combining probabilities

Addition

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

If A and B are mutually exclusive then

$$P(A \text{ or } B) = P(A) + P(B)$$

Multiplication

$$P(A \text{ and } B) = P(A)P(B \mid A)$$

If A and B are independent then

$$P(A \text{ and } B) = P(A)P(B)$$

3 Probability distributions

Discrete distributions	$P(X = r)$	Mean	Variance
Binomial, $X \sim \mathrm{Bin}(n, p)$	${}^{n}\mathrm{C}_{r} p^{r}(1-p)^{n-r}$	np	$np(1-p)$
Poisson, $X \sim \mathrm{Po}(\lambda)$	$\frac{e^{-\lambda}\lambda^{r}}{r!}$	λ	λ

Continuous distributions	$P(X \leq x)$	Mean	Variance
Uniform, $X \sim U(a, b)$	$\dfrac{x-a}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential, $X \sim \mathrm{Exp}(\lambda)$	$1 - e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
Normal, $X \sim N(\mu, \sigma^2)$	$P\left(Z \leq \frac{X-\mu}{\sigma}\right)$, where $Z \sim N(0, 1)$	μ	σ^2

4 Confidence intervals and hypothesis tests

Confidence interval for the population mean $\boldsymbol{\mu}$

σ known: $\bar{x} \pm z \times \sqrt{\sigma^2/n}$;
σ unknown: $\bar{x} \pm t \times \sqrt{s^2/n}$,

where z is a critical value from the standard Normal distribution (i.e. $z = 1.645$, 1.96 and 2.576 for a 90%, 95% and 99% interval (respectively), and t is a critical value from the t distribution with $\nu = n - 1$ degrees of freedom.

Hypothesis test	Test statistic	Other comments
Tests for one mean		
σ known	$z = \frac{\lvert\bar{x}-\mu\rvert}{\sqrt{\sigma^2/n}}$	
σ unknown	$t = \frac{\lvert\bar{x}-\mu\rvert}{\sqrt{s^2/n}}$	$\nu = n - 1$
Tests for two means		
σ_1, σ_2 known	$z = \frac{\lvert\bar{x}_1-\bar{x}_2\rvert}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	
σ_1, σ_2 unknown	$t = \frac{\lvert\bar{x}_1-\bar{x}_2\rvert}{s\times\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$s = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ $\nu = n_1 + n_2 - 2$
Goodness–of–fit tests	$X^2 = \sum \frac{(O-E)^2}{E}$	$\nu =$ (no. of categories after pooling) $-$(no. of estimated parameters) $- 1$
Tests for independence	$X^2 = \sum \frac{(O-E)^2}{E}$	$E = \frac{\text{row total}\times\text{column total}}{\text{overall total}}$ $\nu =$ (no. of rows $-$ 1) $\times$ (no. of columns $-$ 1)

5 Correlation and linear regression

The sample correlation coefficient is given by

$$r = \frac{S_{XY}}{\sqrt{S_{XX} \times S_{YY}}},$$

where

$$\begin{aligned} S_{XY} &= \left(\sum xy\right) - n\bar{x}\bar{y}, \\ S_{XX} &= \left(\sum x^2\right) - n\bar{x}^2, \qquad \text{and} \\ S_{YY} &= \left(\sum y^2\right) - n\bar{y}^2. \end{aligned}$$

The simple linear regression equation is given by

$$Y = \alpha + \beta X + \epsilon,$$

where $\{\epsilon_i\}$ are independent $N\left(0, \sigma^2\right)$ random variables and α and β can be estimated using

$$\begin{aligned} \hat{\beta} &= \frac{S_{XY}}{S_{XX}} \qquad \text{and} \\ \hat{\alpha} &= \bar{y} - \hat{\beta}\bar{x}. \end{aligned}$$

6 Statistical tables

				p		
	One–tailed test	10%	5%	2.5%	1%	0.5%
	Two–tailed test	20%	10%	5%	2%	1%
	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.449
	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
ν	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	∞	1.282	1.645	1.960	2.326	2.576

Table 1: This table contains values of t for which $\Pr(T > t) = p$, where $T \sim t_\nu$. In a two–tailed test, the tabulated values correspond to $\Pr(|T| > t) = p$

				p		
		50%	10%	5%	1%	0.1%
	1	0.45	2.17	3.84	6.63	10.83
	2	1.39	4.61	5.99	9.21	13.82
	3	2.37	6.25	7.82	11.34	16.27
	4	3.36	7.78	9.49	13.28	18.47
	5	4.34	9.24	11.07	15.09	20.52
	6	5.35	10.64	12.59	16.81	22.46
	7	6.35	12.02	14.07	18.48	24.32
ν	8	7.34	13.36	15.51	20.09	26.13
	9	8.34	14.68	16.92	21.67	27.88
	10	9.34	15.99	18.31	23.21	29.59
	12	11.34	18.55	21.03	26.22	32.91
	15	14.34	22.31	25.00	30.58	37.70
	20	19.34	28.41	31.41	37.57	45.32
	25	24.34	34.38	37.65	44.31	52.62
	30	29.34	40.26	43.77	50.89	59.70

Table 2: This table contains values of x for which $\Pr(X^2 > x) = p$, where $X^2 \sim \chi^2_\nu$

z	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0.00
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Table 3: This table contains values of $\Pr(Z < z)$, where $Z \sim N(0,1)$

.

THE END