University of Bath

DEPARTMENT OF ECONOMICS AND INTERNATIONAL DEVELOPMENT

Year 1 : Semester I : EC10003

CORE SKILLS FOR ECONOMISTS: INTRODUCTION TO STATISTICS

Wednesday, 25 January 2006. 1630 – 1830

A University calculator is required for this examination

Additional Materials supplied:
Formulae and Distribution Tables

ANSWER ALL OF SECTION A

AND

ANY TWO QUESTIONS FROM SECTION B

(SECTION A IS WORTH 30% and SECTION B IS WORTH 70%
OF THE FINAL EXAM MARK)

EC10003: Core Skills for Economists: Introduction to Statistics

Read the beginning of each section for instructions specific to that section.

Section A

Instructions

There are 15 questions in this section. You must answer ALL questions.
Each counts 2% toward the overall exam mark.

To obtain credit for a question you must:

i. Clearly state whether you believe the given statement is TRUE or FALSE.
ii. Provide a brief argument for your answer.

This means that if either there is no argument for your answer, or if you do not clearly state whether you believe the statement is TRUE or FALSE, you will receive NO CREDIT.

Questions

1. Any event and its complement are collectively exhaustive, but not necessarily mutually exclusive.

2. When rolling two fair six-sided dice, the probability that the sum of the two rolls is 4, is higher than the probability that the sum is 5.

3. Let the random variable X follow a normal distribution with mean 1 and variance 4. The value of k that satisfies $P(0.5<X<k) = 0.75$, is $k=1.70$

4. The correlation between two random variables X and Y is positive. Then X and Y cannot be independent.

5. Consider drawing two cards without replacement from a deck of 52 cards. The probability that you draw two aces is 1 in 221.

6. The random variable X has the following probability mass function:

$f(X=1) = 0.10$

$f(X=2) = 0.20$

$f(X=3) = 0.30$

$f(X=4) = 0.40$

$f(X) = 0$ otherwise.

Then $P(X > 2.5) = 0.70$

7. During the last 10 years 32% of 18 year olds went to a university. 50% of the 18 year old population were women. The probability that an 18 year old woman went to a university is 34%. Therefore the probability that someone that went to a university (during the past 10 years) is a woman is 63%.

8. When sampling from a normal distribution, the expected value of the sample mean also follows a normal distribution.

9. Consider random sampling size 10 on a random variable X that follows the Poisson distribution with parameter $\lambda = 3$. Then, $E(\overline{X}) = 3$.

10. When sampling from a normal population, the sample mean follows a standard normal distribution.

11. The Mean Squared Error of the sample mean as an estimator of the population mean is always no greater than the population variance.

12. When sampling from a normal population with unknown mean and variance, the expected value of the standardised sample mean is always zero.

13. A 90% confidence interval includes the estimated population parameter with a probability of 10%.

14. The level of significance of a test is equal to the probability of rejecting a hypothesis when false.

15. As the hypothesised value approaches the true value of a parameter, the probability of committing a Type I error increases.

Section B

Instructions

There are 3 questions in this section. You must answer 2 out of the three questions. Each question counts toward 35% of the overall test mark.

The percentages in parentheses indicate the weight of each sub-question within a question. Sub-questions with no indicated weights count equally.

You MUST show all the work necessary to answer a question. If you do not show all calculations, or if you do not state all the logical steps necessary, your answer will be marked down significantly.

Questions

1. (a) Define the random variable X as follows:

X = 1 if when rolling a pair of fair six-sided dice we obtain at least one 3 (i.e. at least one die must be a 3).

X = 0 for any other roll.

(i) Define the pmf of X.

Suppose that you roll the dice twice, hence obtaining a sample of size 2 from this population.

(ii) Derive the sampling distribution of the sample mean.

(iii) Show that the sample mean is an unbiased estimator of the population mean.

(iv) What is the probability that the sample mean is between 0.10 and 0.40?

(75%)

(b) The r.v. X follows a uniform distribution with pdf:

$f(x) = 1/10$ for $11 \leq x \leq 21$

$f(x) = 0$ otherwise.

Calculate the following probabilities:

(i) $P(9 \leq x \leq 14)$

(ii) $P(x > 19)$

(25%)

Formulae

Number of possible outcomes when sampling r objects from n objects:

	Without replacement	With replacement
Order matters	${}_nP_r$	n^r
Order does not matter	${}_nC_r$	${}_{(n+r-1)}C_r$

Where:

${}_nP_r = n! / (n-r)!$

${}_nC_r = n! / [r!\,(n-r)!]$

Bayes' rule:

If events $A_1, A_2, \ldots, A_n$ are mutually exclusive and collectively exhaustive, and A_i is any one of these events, then:

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \ldots + P(A_n)P(B \mid A_n)}$$

Expectations

For a discrete r.v. X with pmf f(x), the expectation or expected value of X is given by:

$$E(X) = \sum_i x_i f(x_i) = \mu_X$$

For a discrete r.v. X with pmf f(x), the variance of X is given by:

$$V(X) = E[(X - \mu_X)^2] = \sum_i (x_i - \mu_X)^2 f(x_i) = \sigma_X^2$$

For a discrete r.v. X with pmf f(x), and a r.v. Y such that Y=g(X) we have:

$$E(Y) = E[g(X)] = \sum_i g(x_i) f(x_i) = \mu_Y$$

For the case where $Y = g(X) = a + bX$, for a, b, constants, we have:

$$\mu_Y = a + b\mu_X$$

$\sigma_Y^2 = b^2 \sigma_X^2$

Parametric distributions

Binomial

pmf: $f(x) = {}_n C_x p^x (1-p)^{(n-x)} = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$

for x=0,1,2, …, n.
f(x) = 0 otherwise.

E(X) = np

V(X) = np(1-p)

Hypergeometric

pmf:

$$f(x) = \frac{({}_S C_x)({}_{N-S} C_{n-x})}{{}_N C_n} = \frac{\frac{S!}{(S-x)!x!} \cdot \frac{(N-S)!}{(N-S-(n-x))!(n-x)!}}{\frac{N!}{(N-n)!n!}}$$

for x=0,1,2, …, n.

f(x) = 0 otherwise.

Where:
N: size of the population.
S: number of observations in the population with attribute of interest.
n: the number of trials.

$$E(X) = n\frac{S}{N}$$

$$V(X) = n\frac{S}{N} \cdot \frac{N-S}{N} \cdot \frac{N-n}{N-1}$$

Poisson

Pmf: $f(x) = e^{-\lambda} \frac{\lambda^x}{x!}$

for x=0, 1, 2, 3, ….

$f(x) = 0$ otherwise

$E(X) = \lambda$

$V(X) = \lambda$

Rectangular

Pdf:

$f(x) = \frac{1}{b-a}$ for $a \le x \le b$

$f(x) = 0$ otherwise

$E(X) = (a+b)/2$

$V(X) = (b-a)^2 / 12$

Exponential

Pdf:

for $\lambda > 0$

$f(x) = \lambda e^{-\lambda x}$ for $x>0$,

$f(x)=0$ otherwise.

$E(X) = \frac{1}{\lambda}$

$V(X) = \frac{1}{\lambda^2}$

Expectations of linear functions of random variables

$Z_1 = a + bX + cY$,

Where a, b, c , are constants. Then:

$E(Z) = a + bE(X) + cE(Y)$

$\mathrm{Var}(Z) = b^2\mathrm{Var}(X) + c^2\mathrm{Var}(Y) + 2bc\mathrm{Cov}(X,Y)$

Let:

$Z_1 = a_1 + b_1X + c_1Y$,

$Z_2 = a_2 + b_2X + c_2Y$,

Where a_i, b_i, c_i , for i= 1,2 are constants. Then

$\mathrm{Cov}(Z_1,Z_2) = b_1b_2\mathrm{Var}(X) + c_1c_2\mathrm{Var}(Y) + (b_1c_2+b_2c_1)\mathrm{Cov}(X,Y)$

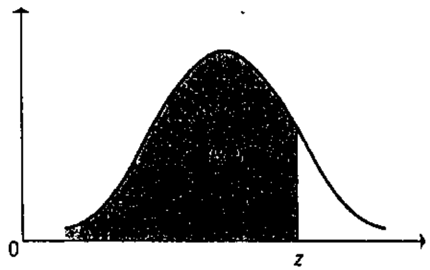
Sample mean

$$\overline{X} = (X_1 + X_2 + X_3 + \ldots + X_n)/n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Sample Variance

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

TABLE 1 Cumulative Distribution Function of the Standard Normal Distribution

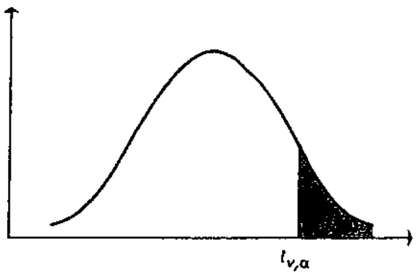


z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)
.00	.5000										
.01	.5040	.31	.6217	.61	.7291	.91	.8186	1.21	.8869	1.51	.9345
.02	.5080	.32	.6255	.62	.7324	.92	.8212	1.22	.8888	1.52	.9357
.03	.5120	.33	.6293	.63	.7357	.93	.8238	1.23	.8907	1.53	.9370
.04	.5160	.34	.6331	.64	.7389	.94	.8264	1.24	.8925	1.54	.9382
.05	.5199	.35	.6368	.65	.7422	.95	.8289	1.25	.8944	1.55	.9394
.06	.5239	.36	.6406	.66	.7454	.96	.8315	1.26	.8962	1.56	.9406
.07	.5279	.37	.6443	.67	.7486	.97	.8340	1.27	.8980	1.57	.9418
.08	.5319	.38	.6480	.68	.7517	.98	.8365	1.28	.8997	1.58	.9429
.09	.5359	.39	.6517	.69	.7549	.99	.8389	1.29	.9015	1.59	.9441
.10	.5398	.40	.6554	.70	.7580	1.00	.8413	1.30	.9032	1.60	.9452
.11	.5438	.41	.6591	.71	.7611	1.01	.8438	1.31	.9049	1.61	.9463
.12	.5478	.42	.6628	.72	.7642	1.02	.8461	1.32	.9066	1.62	.9474
.13	.5517	.43	.6664	.73	.7673	1.03	.8485	1.33	.9082	1.63	.9484
.14	.5557	.44	.6700	.74	.7704	1.04	.8508	1.34	.9099	1.64	.9495
.15	.5596	.45	.6736	.75	.7734	1.05	.8531	1.35	.9115	1.65	.9505
.16	.5636	.46	.6772	.76	.7764	1.06	.8554	1.36	.9131	1.66	.9515
.17	.5675	.47	.6803	.77	.7794	1.07	.8577	1.37	.9147	.167	.9525
.18	.5714	.48	.6844	.78	.7823	1.08	.8599	1.38	.9162	1.68	.9535
.19	.5753	.49	.6879	.79	.7852	1.09	.8621	1.39	.9177	1.69	.9545
.20	.5793	.50	.6915	.80	.7881	1.10	.8643	1.40	.9192	1.70	.9554
.21	.5832	.51	.6950	.81	.7910	1.11	.8665	1.41	.9207	1.71	.9564
.22	.5871	.52	.6985	.82	.7939	1.12	.8686	1.42	.9222	1.72	.9573
.23	.5910	.53	.7019	.83	.7967	1.13	.8708	1.43	.9236	1.73	.9582
.24	.5948	.54	.7054	.84	.7995	1.14	.8729	1.44	.9251	1.74	.9591
.25	.5987	.55	.7088	.85	.8023	1.15	.8749	1.45	.9265	1.75	.9599
.26	.6026	.56	.7123	.86	.8051	1.16	.8770	1.46	.9279	1.76	.9608
.27	.6064	.57	.7157	.87	.8078	1.17	.8790	1.47	.9292	1.77	.9616
.28	.6103	.58	.7190	.88	.8106	1.18	.8810	1.48	.9306	1.78	.9625
.29	.6141	.59	.7224	.89	.8133	1.19	.8830	1.49	.9319	1.79	.9633
.30	.6179	.60	.7257	.90	.8159	1.20	.8849	1.50	.9332	1.80	.9641

[illegible] 1.20 .8849 | 1.50 .9332 | 1.80 .9641

z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)	z	F(z)
1.81	.9649	2.21	.9864	2.61	.9955	3.01	.9987	3.41	.9997	3.81	.9999
1.82	.9656	2.22	.9868	2.62	.9956	3.02	.9987	3.42	.9997	3.82	.9999
1.83	.9664	2.23	.9871	2.63	.9957	3.03	.9988	3.43	.9997	3.83	.9999
1.84	.9671	2.24	.9875	2.64	.9959	3.04	.9988	3.44	.9997	3.84	.9999
1.85	.9678	2.25	.9878	2.65	.9960	3.05	.9989	3.45	.9997	3.85	.9999
1.86	.9686	2.26	.9881	2.66	.9961	3.06	.9989	3.46	.9997	3.86	.9999
1.87	.9693	2.27	.9884	2.67	.9962	3.07	.9989	3.47	.9997	3.87	.9999
1.88	.9699	2.28	.9887	2.68	.9963	3.08	.9990	3.48	.9997	3.88	.9999
1.89	.9706	2.29	.9890	2.69	.9964	3.09	.9990	3.49	.9998	3.89	1.0000
1.90	.9713	2.30	.9893	2.70	.9965	3.10	.9990	3.50	.9998	3.90	1.0000
1.91	.9719	2.31	.9896	2.71	.9966	3.11	.9991	3.51	.9998	3.91	1.0000
1.92	.9726	2.32	.9898	2.72	.9967	3.12	.9991	3.52	.9998	3.92	1.0000
1.93	.9732	2.33	.9901	2.73	.9968	3.13	.9991	3.53	.9998	3.93	1.0000
1.94	.9738	2.34	.9904	2.74	.9969	3.14	.9992	3.54	.9998	3.94	1.0000
1.95	.9744	2.35	.9906	2.75	.9970	3.15	.9992	3.55	.9998	3.95	1.0000
1.96	.9750	2.36	.9909	2.76	.9971	3.16	.9992	3.56	.9998	3.96	1.0000
1.97	.9756	2.37	.9911	2.77	.9972	3.17	.9992	3.57	.9998	3.97	1.0000
1.98	.9761	2.38	.9913	2.78	.9973	3.18	.9993	3.58	.9998	3.98	1.0000
1.99	.9767	2.39	.9916	2.79	.9974	3.19	.9993	3.59	.9998	3.99	1.0000
2.00	.9772	2.40	.9918	2.80	.9974	3.20	.9993	3.60	.9998		
2.01	.9778	2.41	.9920	2.81	.9975	3.21	.9993	3.61	.9998		
2.02	.9783	2.42	.9922	2.82	.9976	3.22	.9994	3.62	.9999		
2.03	.9788	2.43	.9925	2.83	.9977	3.23	.9994	3.63	.9999		
2.04	.9793	2.44	.9927	2.84	.9977	3.24	.9994	3.64	.9999		
2.05	.9798	2.45	.9929	2.85	.9978	3.25	.9994	3.65	.9999		
2.06	.9803	2.46	.9931	2.86	.9979	3.26	.9994	3.66	.9999		
2.07	.9808	2.47	.9932	2.87	.9979	3.27	.9995	3.67	.9999		
2.08	.9812	2.48	.9934	2.88	.9980	3.28	.9995	3.68	.9999		
2.09	.9817	2.49	.9936	2.89	.9981	3.29	.9995	3.69	.9999		
2.10	.9821	2.50	.9938	2.90	.9981	3.30	.9995	3.70	.9999		
2.11	.9826	2.51	.9940	2.91	.9982	3.31	.9995	3.71	.9999		
2.12	.0830	2.52	.9941	2.92	.9982	3.32	.9996	3.72	.9999		
2.13	.9834	2.53	.9943	2.93	.9983	3.33	.9996	3.73	.9999		
2.14	.9838	2.54	.9945	2.94	.9984	3.34	.9996	3.74	.9999		
2.15	.9842	2.55	.9946	2.95	.9984	3.35	.9996	3.75	.9999		
2.16	.9846	2.56	.9948	2.96	.9985	3.36	.9996	3.76	.9999		
2.17	.9850	2.57	.9949	2.97	.9985	3.37	.9996	3.77	.9999		
2.18	.9854	2.58	.9951	2.98	.9986	3.38	.9996	3.78	.9999		
2.19	.9857	2.59	.9952	2.99	.9986	3.39	.9997	3.79	.9999		
2.20	.9861	2.60	.9953	3.00	.9986	3.40	.9997	3.80	.9999		

TABLE 8 Cutoff Points for the Student's *t* Distribution



For selected probabilities, α, the table shows the values $t_{v,\alpha}$ such that $P(t_v > t_{v,\alpha}) = \alpha$, where t_v is a Student's *t* random variable with v degrees of freedom. For example, the probability is .10 that a Student's *t* random variable with 10 degrees of freedom exceeds 1.372.

v	α				
	0.100	**0.050**	**0.025**	**0.010**	**0.005**
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	3.499
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
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
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
∞	1.282	1.645	1.960	2.326	2.576

2. (a) In the country of Bathland, university graduates are paid more that non-graduates. In that country the amount of 100 zgups is a substantial sum of money. The joint frequency of the population by being a university graduate and earning more than 100 zgups annually is given below.

	100 or more	Less than 100
Graduate	0.21	0.11
Non-graduate	0.23	0.45

Let the random variable X be such that:

X = 0 if a person is not a university graduate
X = 1 if a person is a university graduate

And the random variable Y:

Y = 0 if a person earns less than 100 zgups
Y = 1 if a person earns at least 100 zgups

(i) Specify the joint pmf of X and Y.

(ii) Specify the marginal pmfs of X and Y.

(iii) Find all the distinct conditional pmf's of Y, conditional upon X.

(iv) Find the conditional expectation function of Y, conditional upon X.

(v) Find the covariance between X and Y.

(vi) Let Z = XY. Find E(Z).

(75%)

(b) You have a random sample size 6 from a normal population with the following observations:

67, 85, 110, 119, 132 and 193.

(i) Test the hypothesis that the sample mean is equal to 120 at the 0.01 significance level.

(ii) How would your answer change if you had a sample size 12, with two observations having each of the above values?

(25%)

3. (a) For year 2003, the logarithm of annual individual earnings followed a normal distribution with mean 9.711 and variance 75.682 (all logarithms are taken to be natural logarithms). For year 2002, the logarithm of annual individual earnings followed a normal distribution with mean 9.702 and variance 75.093. The correlation between the two log earnings was 0.73. Let Y denote an individual's average log earnings over the two years.

(i) If you randomly draw one individual, what is the probability that Y is more than 10.223?

(ii) If you draw a random sample of 3 individuals, what is the probability that the sample mean of Y is more than 10.223?

(50%)

(b) You randomly draw five cards out of a deck of 52 cards.

(i) What is the probability that you draw exactly three Kings?

(ii)What is the probability that you draw exactly two cards of the same (arbitrary) denomination?

(iii) What is the probability that all five cards are of the same (arbitrary) colour?

(50%)