ADVANCED PROBABILITY AND STATISTICS (MATH 4281)

Time Allotted : 3 hrs **Full Marks : 70**

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.

Candidates are required to give answer in their own words as far as practicable.

Group – A
(Multiple Choice Type Questions)

1. Choose the correct alternative for the following: **10 × 1 = 10**

(i) The regression lines between two random variables X and Y are given by 3x + y = 10 and 3x + 4y = 12. Let $u = 3x + 2,\ v = -2y + 5$, then the correlation coefficient between $u$and$v$is:
(a) -0.5. (b) 0.5. (c) -1. (d) 1.

(ii) If X represent the number of six, when a fair die is rolled 10 times. The moment generating function $M_X(t)$of X is:
(a) $e^{\frac{1}{6}(e^t-1)}$ (b) $e^{\frac{t^2}{2}}$ (c) $\left(\frac{1}{6}e^t+\frac{5}{6}\right)^{10}$ (d) $\left(\frac{5}{6}e^t+\frac{1}{6}\right)^{10}$

(iii) For normal distribution, the coefficient of skewness $(\gamma_1)$and kurtosis $(\gamma_2)$are:
(a) $\gamma_1=1, \gamma_2=0$. (b) $\gamma_1=0, \gamma_2=3$.
(c) $\gamma_1=0, \gamma_2=0$. (d) $\gamma_1=3, \gamma_2=-3$.

(iv) Let the spectrum of a random variable X is given by $\{-10, -9, -8, \cdots\}$. It is given that $E(X) = -6$ then the upper bound of $P(X \geq -2)$is:
(a) -0.5. (b) 0.5. (c) -1 (d) 1.

(v) If X and Y are independent random variables, then from the following table $P(X = 1|\ Y = 5)$ is:

Y \ X	0	5
1	0.3	0.1
5	0.2	0.4

(a) 0.4. (b) 0.6. (c) 0. (d) 0.5.

(vi) If X is random variable, then which **one** of the following is **correct**?
(a) $E|X| \leq |E(X)|$ (b) $E|X| \geq |E(X)|$
(c) $E|X| = \frac{|E(X)|}{2}$ (d) $E|X| \neq |E(X)|$

(vii) The probability density function of a random variable X is $f(x) = \begin{Bmatrix} 3x^2, 0 < x < 1 \\ 0, elsewhere \end{Bmatrix}$.
If $g(y)$ is the probability density function of $Y = 2X + 3$, then $g(y)$ is:
(a) $\frac{(y-3)^2}{16}$. (b) $\frac{y-3}{4}$. (c) $\frac{y}{4}$. (d) $6\ (y-3)^2$.

(viii) If X be the test statistic and (a, b) is the region of acceptance corresponding to 3% level of significance, then $P(a \le X \le b) = ?$
(a) 0.9. (b) 0.97. (c) 0.99. (d) 0.03.

(ix) If$\lim_{n\to\infty} P\left\{\left|\frac{1}{\overline{X}+2} - p\right| > \epsilon\right\} = 0$ hold for arbitrary $\epsilon > 0$, then
(a) $\overline{X}$ is consistent estimator of p.
(b) $\overline{X} + 2$ is consistent estimator of p.
(c) $\frac{1}{\overline{X}}$ is consistent estimator of p.
(d) $\frac{1}{\overline{X}+2}$ is consistent estimator of p.

(x) Which **one** of the following is **NOT** a possible pair of values of $(E(X), E(X^2))$
(a) $\left(\frac{2}{3}, 1\right)$. (b) $(5, 9)$. (c) $\left(\frac{1}{2}, \frac{3}{4}\right)$. (d) $(2.5, 5)$.

Group – B

2. (a) A bag contains 5 white, 7 green, 12 red and 14 black balls. The balls are indistinguishable from each other except by colour. A ball is drawn and replaced after its colour has been noted, on ten occasions. If any ball is likely to be drawn as any other, find the probability that of the ten balls seen 3 will be white, 3 green, 2 red and 2 black.

(b) Assume that 5% of the population over 65 years old has Alzheimer's disease. Suppose a random sample of 9600 people over 65 is taken. Using normal approximationto binomial distribution, find the probability thatmore than 400 and fewer than 500 of them have the disease.

(c) An insurance company has discovered that only 0.1% of the population is involved in a certain type of accident each year. If its 10,000 policy holders were randomly selected from the population, what is the probability that not more than 5 of its clients are involved in such an accident next year?

3 + 5 + 4 = 12

3. (a) If the density function of a continuous random variable is given by
$$f_X(x) = \begin{cases} \frac{2}{9}(x+1), & -1 < x < 2 \\ 0, & otherwise \end{cases}.$$
Find the density function of $Y = \frac{1}{2}(X + |X|)$.

(b) The amount of time that a watch will run without having to be reset is a random variable having an exponential distribution with mean 120 days. Find the probability that such a watch will
(i) have to be reset in less than 24 days, and

(ii) not have to be reset in at least 180 days.

6 + 6 = 12

Group – C

4. (a) Suppose that an examination contains 99 questions arranged in a sequence from the easiest to the most difficult. Supposethat the probability that a particular student will answer the first questioncorrectly is 0.99; the probability that he will answer the second question correctly is 0.98; and in general, the probability that he will answer the i^{th} questioncorrectly is
$1 - \frac{\mathrm{i}}{100}$ for $\mathrm{i} = 1,2,\cdots,99$.
It is assumed that all questions will be answered independently and that the student must answer at least 60 questions, correctly to pass the examination. What is the probability that the student will pass? **(Use Central limit theorem)**

(b) The joint probability density function of the random variables X and Y is given by $f(x,y) = \begin{cases} 2, 0 < x < 1, 0 < y < x \\ 0 \quad ,elsewhere \end{cases}$.
Find the marginal and conditional density functions. Also compute
$P\left(\frac{1}{4} < X < \frac{3}{4} \mid Y = \frac{1}{2}\right)$.

(c) If the random variable X is uniformly distributed over $(-\frac{\pi}{2}, \frac{\pi}{2})$, find the probability density function of $Y = \sin X$, using characteristic function.

4 + 5 + 3 = 12

5. (a) Use Chebyshev's inequality to find how many times a fair die must be tossed in order that the probability that the ratio of the number of six to the number of tosses will lie between 0.4 and 0.6 will be at least 0.9.

(b) (i) Let X be a random variable with mean μ and variance σ^2, and let $M_X(t)$ denotes the moment generating function of X for $-\infty < t < \infty$. Let c be a given positive constant, and let Y be a random variable for which the moment generating function is
$M_Y(t) = e^{c(M_X(t)-1)}$ for $-\infty < t < \infty$
Find expressions for the mean and the variance of Y in terms of the meanand the variance of X.

(ii) Let X_1, X_2, X_3 be three mutually independent random variables from a discrete distribution with probability function

$$P(X = x) = \begin{cases} \frac{1}{4}, & x = -1 \\ \frac{1}{2}, & x = 0 \\ \frac{1}{4}, & x = 1 \\ 0, & elsewhere \end{cases}$$

Determine the moment generating function, $M(t)$ of $Y = X_1 X_2 X_3$.

5 + (2 + 5) = 12

Group – D

6. (a) Suppose that eight specimens of a certain type of alloy were produced at different temperatures, and that the durability of each specimen was then observed. The observed values are as given in following Table, where y denotesthe temperature (in coded units) at which specimen was produced andxdenotes the durability (in coded units) of that specimen.

x	44	41	43	42	44	42	43	42
y	0.5	1	1.5	2	2.5	3	3.5	4

Fit a parabola of the form $x = a + by + cy^2$ to these values by the method of least squares.

(b) Find the moment-generating function of the random variable X which follows Normal distribution with parameter μ and σ^2. Hence use it to find the mean and variance of random variable$Y = \frac{X-3}{\sqrt{2}}$.

6 + 6 = 12

7. (a) Calculate skewness and kurtosis for the following distribution and comment on the nature of the distribution.

x	0	1	2	3	4	5	6	7	8
f	1	8	28	56	70	56	28	8	1

(b) The following marks obtained by a class of Engineering students:

Paper-I(x)	80	45	55	56	58	60	65	68	70	75	85
Paper-II(y)	81	56	50	48	60	62	64	65	70	74	90

Find (i) correlation coefficient of x and y,
(ii) the regression line of 'y on x' and 'x on y',
and (iii) analyse the nature of association between x and y.

6 + 6 = 12

Group – E

8. (a) A Sample of 900 members is found to have a mean of 3.4 cms. Can it be reasonably regarded as a simple sample from a large population with mean 3.2 cms and standard deviation 2.3 cms?

(b) If $\{x_1, x_2, x_3, \ldots.., x_n\}$ are random observations of avariable X, taking the value 1 with probability θ and the value 0 with probability $1-\theta$. Show that $\frac{T(T-1)}{n(n-1)}$ is an unbiased estimate of θ^2, where $T = \sum_{i=1}^{n} x_i$.

(c) Let p denote the probability of getting head when a given coin is tossed once. Suppose that the null hypothesis $H_0: p = 0.5$ is rejected in favour of $H_1: p = 0.7$ if10 trials results 8 or more heads. Calculate the probabilities of Type-I error and Type-II error.

3 + 5 + 4 = 12

9. (a) A random variable X can take all non-negative integral values and $P(X = i) = p(1-p)^i$, $i = 0, 1, 2, \ldots.$, where, $p\ (0 < p < 1)$ is a parameter. Find the maximum likelihood estimate of p on the basis of a random sample $\{x_1, x_2, x_3, \ldots., x_n\}$ of size n drawn from the population of X.

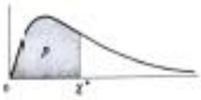
(b) 12 Samples were taken from a large stock of nuts and their weights were recorded (in *grams*) as 3.86, 3.50, 4.12, 3.67, 4.08, 3.61, 3.79, 4.01, 4.05, 3.91, 3.97 and 3.72. Find 98% confidence interval for the standard deviation of the population of weights assuming the population to be normal.

6 + 6 = 12

Department & Section	Submission Link
ECE, CSE, IT	https://classroom.google.com/c/Mjk4MTA5MzQxOTYx/a/MzQ0NDg4NDcwNDIx/details

Table A-3 The chi-square distribution

The entry in row k (degrees of freedom) under column heading p (probability) is the value χ^* for which $P(0 \le \chi^2 \le \chi^*) = p$.

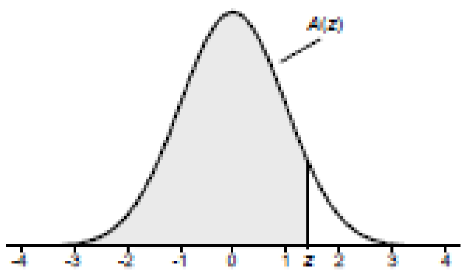


k \ p	0.005	0.01	0.025	0.05	0.10	0.25	0.50	0.75	0.90	0.95	0.975	0.99	0.995	0.999
1	.0000	.0002	.0010	.0039	.0158	.102	.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	.0100	.0201	.0506	.103	.211	.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	.0717	.115	.216	.352	.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	.207	.297	.484	.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	.412	.554	.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	.676	.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	73.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	86.7
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	99.6
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100	104	112
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	102	107	112	116	125
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	108	113	118	124	128	137
100	67.3	70.1	74.2	77.9	82.4	90.1	99.3	109	118	124	130	136	140	149

Source: E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1 (1966), Table 8, pages 137 and 138, by permission.

Table A.1

Cumulative Standardized Normal Distribution



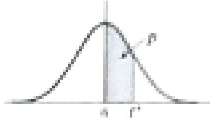
$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

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
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Table A-2 The t distribution

The entry in row k (degrees of freedom) under column heading p (probability) is the value t^* for which $P(0 \le t \le t^*) = p$.



k \ p	0.05	0.1	0.2	0.25	0.3	0.4	0.45	0.475	0.49	0.495
1	.158	.325	.727	1.000	1.376	3.08	6.31	12.71	31.82	63.66
2	.142	.289	.617	.816	1.061	1.89	2.92	4.30	6.96	9.92
3	.137	.277	.584	.765	.978	1.64	2.35	3.18	4.54	5.84
4	.134	.271	.569	.741	.941	1.53	2.13	2.78	3.75	4.60
5	.132	.267	.559	.727	.920	1.48	2.02	2.57	3.36	4.03
6	.131	.265	.553	.718	.906	1.44	1.94	2.45	3.14	3.71
7	.130	.263	.549	.711	.896	1.42	1.90	2.36	3.00	3.50
8	.130	.262	.546	.706	.889	1.40	1.86	2.31	2.90	3.36
9	.129	.261	.543	.703	.883	1.38	1.83	2.26	2.82	3.25
10	.129	.260	.542	.700	.879	1.37	1.81	2.23	2.76	3.17
11	.129	.260	.540	.697	.876	1.36	1.80	2.20	2.72	3.11
12	.128	.259	.539	.695	.873	1.36	1.78	2.18	2.68	3.06
13	.128	.259	.538	.694	.870	1.35	1.77	2.16	2.65	3.01
14	.128	.258	.537	.692	.868	1.34	1.76	2.14	2.62	2.98
15	.128	.258	.536	.691	.866	1.34	1.75	2.13	2.60	2.95
16	.128	.258	.535	.690	.865	1.34	1.75	2.12	2.58	2.92
17	.128	.257	.534	.689	.863	1.33	1.74	2.11	2.57	2.90
18	.127	.257	.534	.688	.862	1.33	1.73	2.10	2.55	2.88
19	.127	.257	.533	.688	.861	1.33	1.73	2.09	2.54	2.86
20	.127	.257	.533	.687	.860	1.32	1.72	2.09	2.53	2.84
21	.127	.257	.532	.686	.859	1.32	1.72	2.08	2.52	2.83
22	.127	.256	.532	.686	.858	1.32	1.72	2.07	2.51	2.82
23	.127	.256	.532	.685	.858	1.32	1.71	2.07	2.50	2.81
24	.127	.256	.531	.685	.857	1.32	1.71	2.06	2.49	2.80
25	.127	.256	.531	.684	.856	1.32	1.71	2.06	2.48	2.79
26	.127	.256	.531	.684	.856	1.32	1.71	2.06	2.48	2.78
27	.127	.256	.531	.684	.855	1.31	1.70	2.05	2.47	2.77
28	.127	.256	.530	.683	.855	1.31	1.70	2.05	2.47	2.76
29	.127	.256	.530	.683	.854	1.31	1.70	2.04	2.46	2.76
30	.127	.256	.530	.683	.854	1.31	1.70	2.04	2.46	2.75
40	.126	.255	.529	.681	.851	1.30	1.68	2.02	2.42	2.70
60	.126	.254	.527	.679	.848	1.30	1.67	2.00	2.39	2.66
120	.126	.254	.526	.677	.845	1.29	1.66	1.98	2.36	2.62
∞	.126	.253	.524	.674	.842	1.28	1.645	1.96	2.33	2.58

Source: R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, published by Longman Group Ltd., London (previously published by Oliver and Boyd, Edinburgh), and by permission of the authors and publishers.