NUI Galway
OÉ Gaillimh

Semester I Examinations 2018/2019

Exam Code(s)	1OA2, 1OA3, 1BY1, 1EV1, 1EM1, 2BA1, 2BCS1, 2BME1, 2BCM1, 2BCW1, 2BCT1, 2BPT1, 2BS1, 2EH1, 2MR1, 3BS9, 3MR3
Exam(s)	First Science, Second Arts, Second Science and Third Science
Module(s)	Introduction to Statistical Data and Probability
Module Code(s)	ST237
Paper No	1
Repeat Paper	
External Examiner	Prof. S. Wilson
Internal Examiner	D. Roshan
	Prof. J. Newell

Instructions:

Answer the 10 Questions in Section A (20 marks)
and
any 4 Questions from Section B (20 marks each).

Duration	2 Hours
No. of Pages	12 Pages
School	School of Mathematics, Statistics and Applied Mathematics

Requirements:

Release to Library:	Yes
Statistical Tables/ Log Tables	Relevant distribution tables are attached to this paper The *New Formulae & Tables* are optional.
Other Materials	A calculator is allowed (non-programmable and not capable of storing text)

Section A - Compulsory Questions (20 marks)

A maximum of two marks will be awarded for each part.

A1. Which of the following measures of central tendency and spread **is affected** by extreme values:

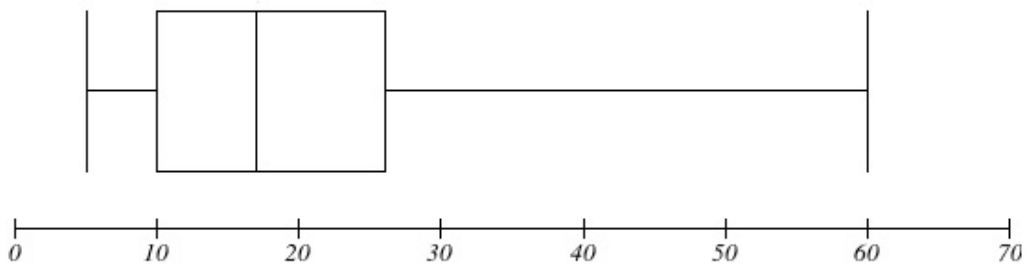
(a) Range

(b) Mean

(c) Standard Deviation

(d) all of the above

A2. The following boxplot shows a sample of data with a median at 18.



Will the mean of the data be

(a) less than the median;

(b) approximately equal to the median;

(c) larger than the median?

A3. The waiting time in minute from time to entry to time at which service begins was recorded for each of a random sample of 14 customers entering a bank and yielded the following data:

30	8	9	39	33	27	14
18	19	10	34	37	23	7

Calculate the median waiting time.

A4. What is the relationship between **standard deviation** and **variance**?
Why is it usually preferable to quote the standard deviation as a measure of spread?

A5. In how many ways can a president, a treasurer and a secretary be chosen from among 7 candidates?

A6. If two events A and B are **mutually exclusive** and have probabilities $P(A) = 0.3$ and $P(B) = 0.6$. Find $P(\overline{A \cup B})$.

A7. If two events A and B are **independent** and if $P(A) = 0.3$ and $P(B) = 0.6$, find $P(\overline{A} \cap \overline{B})$.

A8. In two rolls of a fair die, what is the probability of **not getting a double**?

A9. A random variable X has the following probability distribution

x	-2	0	2	4	6
$P(X = x)$	0.30	0.20	0.1	0.25	0.15

Find E $[X]$, the expected (mean) value of X.

A10. For a random variable X with variance 25, find the standard deviation of $3X - 5$.

Section B — Answer 4 Questions

B1. (a) An insurance company has collected the following data on the number of car thefts per day in a large city for a period of 21 days:

52	55	53	33	57	61	58	65	48	57	44
76	59	54	69	64	97	57	80	70	59	

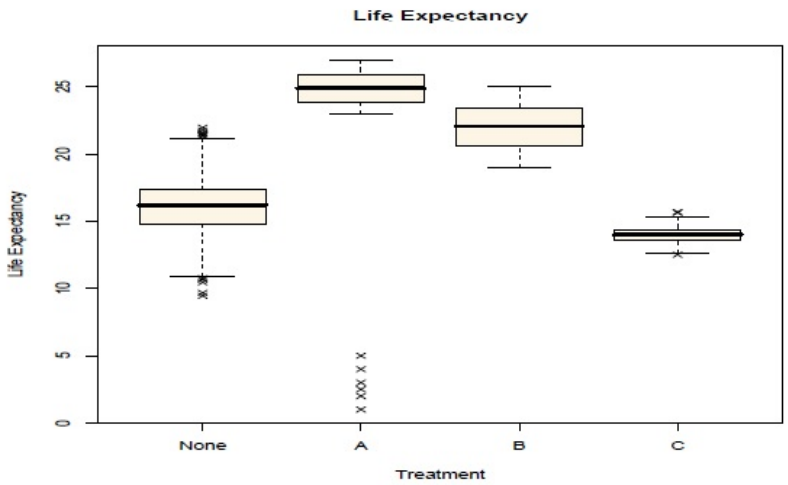
i) Construct a stem-and-leaf plot using intervals for the stems of width 10. (4)

ii) Find the five number summary. (5)

iii) Draw a boxplot for these data (you do not need to use graph paper for this, but you may do so if you wish). (3)

iv) Calculate the sample mean and sample standard deviation. (4)

(b) The median life expectancy from diagnosis of a certain type of cancer is 16 months. Three drugs to treat this cancer are each tested on $1,000$ patients, and the time (in months) from diagnosis to death (i.e. life expectancy) is recorded for each patient. The resultant data are illustrated by the multiple boxplot presented below. The boxplot labeled "None" illustrates the times-to-death of patients that received no treatment.
Briefly comment upon the effectiveness of the drugs A, B and C with regard to increasing life expectancy (or otherwise). (4)



B2. (a) i) How many different ways are there of writing down (in a sequence) the letters in (3)

PASTEXAMPAPERS

ii) What is the probability that a random rearrangement of these letters contains the string '**STAR**'? (2)

(b) Consider an urn containing 12 balls, of which 8 are white. A sample of size 2 is to be drawn without replacement. What is the probability that

i) the first ball is white? (1)

ii) the second ball drawn will be white given that the first drawn ball was white? (2)

iii) the second one is white (2)

(c) A statistics professor classifies his students according to their grade and their gender. It was found that approximately 40% are male and 38% have grade under 40. For males the (*conditional*) probability of having grade under 40 is 0.375.

i) Find the joint probability of being male and having a grade under 40. (2)

ii) Hence, complete a table of joint and marginal probabilities as begun below. (2)

		Garde		
		Under 40	Above 40	Total
Gender	Male			0.4
	Female			
	Total	0.38		1.0

iii) Is having low grade (i.e. under 40) independent of gender? (1)

(d) A total of 46% of the voters in a certain city classify themselves as Independents, whereas 30% classify themselves as Liberals and 24% say that they are Conservatives. In a recent local election, 35% of the Independents, 62% of the Liberals, and 58% of the Conservatives voted.

i) What fraction of voters participated in the local election? (3)

ii) A voter is chosen at random. Given that this person voted in the local election, what is the probability that he or she is an Independent? (2)

B3. (a) In a large restaurant an average of 3 out of every 5 customers ask for water with their meal. A random sample of 10 customers is selected. Find the probability that:

i) Exactly 6 ask for water with their meal. (2)

ii) Less than 9 ask for water with their meal. (2)

If a second sample of 10 customers is selected, what is the probability that in one of the samples only 8 will ask for water while all requesting water in another sample. (3)

Note: relevant tables of the binomial distribution are given at the back of this paper.

(b) In constructing a gas pipeline the number of faulty welds is Poisson distributed with a mean of 1 per kilometer. Find the probability that

i) there are no faulty welds in a 1 kilometer section; (2)

ii) there are at least two faulty welds in a 1 kilometer section; (2)

iii) there are no faulty welds over a 5 kilometer section (assuming the rate is constant over the kilometers); (2)

(c) The weights of bags of red gravel may be modelled by a normal distribution with mean 25.8kg and standard deviation 0.5kg.

i) What is the probability that a randomly selected bag of red gravel will weigh
   - less than 25kg? (2)
   - between 25.5 and 26.5? (2)

ii) Determine, to two decimal places, the weight exceeded by 75% of bags? (3)

B4. The joint probability mass function of two random variables X and Y is given by the following table:

		Y		
		1	2	3
	-2	0.1	0.25	0.1
X	0	0.2	0.05	0
	2	0	0.15	0.15

(a) Find the marginal probability functions for X and Y. (3)

(b) What is the conditional probability that $Y = 3$ given that $X = 2$? (2)

(c) Calculate the mean and variance of each of X and Y. (6)

(d) Calculate the covariance between X and Y. (3)

(e) Find the correlation between X and Y. (2)

(f) Are X and Y independent? Justify your answer. (2)

(g) What is $P(X + Y > 3)$? (2)

B5. (a) A shopkeeper believes that 15% of customers buy an item from the display next to the till. He will take a random sample of 100 customers.

i) What is an appropriate binomial probability model for the number of people, in the sample, who will buy an item from the display next to the till? (3)

ii) Use a normal approximation to the binomial distribution to calculate the probability that more than 22 people will buy an item from the display next to the till. (4)

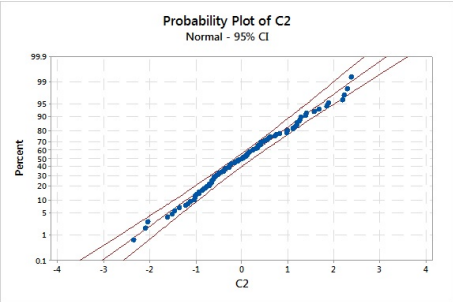
iii) What are the two assumptions of the binomial model? Are they reasonable here? (3)

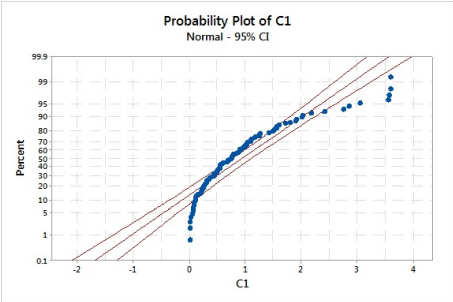
(b) The normal distribution plays an important role in statistics. Write a very brief comment/answer to each of the following aspects:

i) the relationship between the mean and the median of the normal distribution; (2)

ii) the $68 - 95 - 99.7$ rule; (3)

iii) the use of normal probability plots, with reference to the two examples below; (3)





iv) the Central Limit Theorem. (2)

Formulæ

- Sample Mean

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

- Sample Standard Deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1}}$$

- Binomial Distribution

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, \ldots, n$$

$$\mathsf{E}[X] = np \qquad \mathsf{Var}(X) = np(1-p)$$

- Poisson Distribution

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, \ldots$$

$$\mathsf{E}[X] = \lambda \qquad \mathsf{Var}(X) = \lambda$$

- Variances, Covariances and Correlation

$$\mathsf{Var}(X) = \mathsf{E}\left[X^2\right] - (\mathsf{E}[X])^2$$

$$\mathsf{cov}[X, Y] = \mathsf{E}[XY] - \mathsf{E}[X]\,\mathsf{E}[Y]$$

$$\mathsf{Correlation}(X, Y) = \frac{\mathsf{cov}[X, Y]}{\sqrt{\mathsf{Var}(X)\,\mathsf{Var}(Y)}}$$

Binomial Table

Table C Binomial probabilities (p.4 of 6)

Entry is $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

		p								
n	k	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	0	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7			.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8				.0001	.0004	.0013	.0035	.0083	.0176
	9						.0001	.0003	.0008	.0020
10	0	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7		.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8			.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9					.0001	.0005	.0016	.0042	.0098
	10							.0001	.0003	.0010
12	0	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
	7		.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8		.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9			.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10					.0002	.0008	.0025	.0068	.0161
	11						.0001	.0003	.0010	.0029
	12								.0001	.0002
15	0	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	
	1	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417
	5	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8		.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9		.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10			.0001	.0007	.0030	.0096	.0245	.0515	.0916
	11				.0001	.0006	.0024	.0074	.0191	.0417
	12					.0001	.0004	.0016	.0052	.0139
	13						.0001	.0003	.0010	.0032
	14								.0001	.0005
	15									

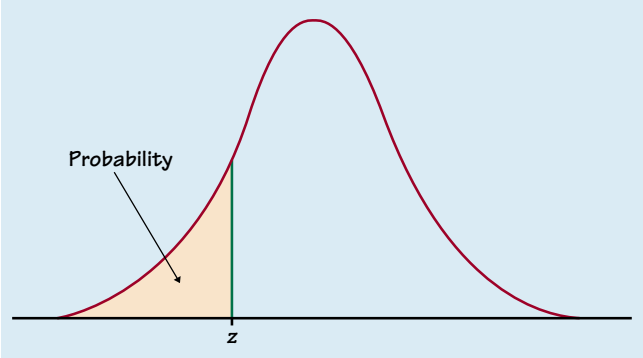


Table entry for z is the probability lying below z.

TABLE A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

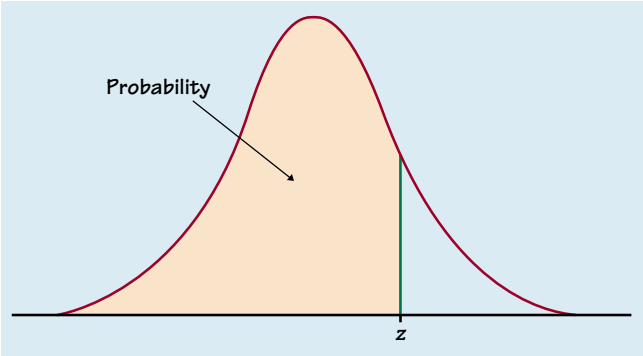


Table entry for z is the probability lying below z.

TABLE A Standard normal probabilities (*Continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998