THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

This exam paper must not be removed from the venue

Venue ____________________

Seat Number ________

Student Number |__|__|__|__|__|__|__|__|

Family Name _____________________

First Name _____________________

School of Mathematics & Physics

EXAMINATION EXAMPLE 2

Semester One Final Examinations, 2016

STAT2201 Analysis of Engineering and Scientific Data

This paper is for St Lucia Campus students.

Examination Duration: 120 minutes

Reading Time: 10 minutes

Exam Conditions:

This is a Central Examination

This is a Closed Book Examination - no materials permitted

During reading time - write only on the rough paper provided

This examination paper will be released to the Library

Materials Permitted In The Exam Venue:

(No electronic aids are permitted e.g. laptops, phones)

Calculators - Any calculator permitted - unrestricted

Materials To Be Supplied To Students:

none

Instructions To Students:

Additional exam materials (eg. answer booklets, rough paper) will be provided upon request.

For Examiner Use Only

Question	Mark
1a	
1b	
1c	
1d	
2a	
2b	
2c	
2d	
3a	
3b	
3c	
3d	
4a	
4b	
4c	
4d	

Total ________

Instructions

The exams consists of 4 questions, 1-4. Each question has four items, a-d.

Within each question:

Item (a) carries a weight of 8 marks.

Item (b) carries a weight of 7 marks.

Item (c) carries a weight of 6 marks.

Item (d) carries a weight of 4 marks.

The total marks in the exam are 100.

Answer ALL questions in the spaces provided. If more space is required, use the back of the PREVIOUS page.

Show all your working and include sketches where appropriate.

Work written in the Formulae and Tables section will NOT be marked.

Question 1

You are presented with a sample composed of $n = 100$ observations $y = \{y_1, y_2, \ldots, y_{100}\}$.

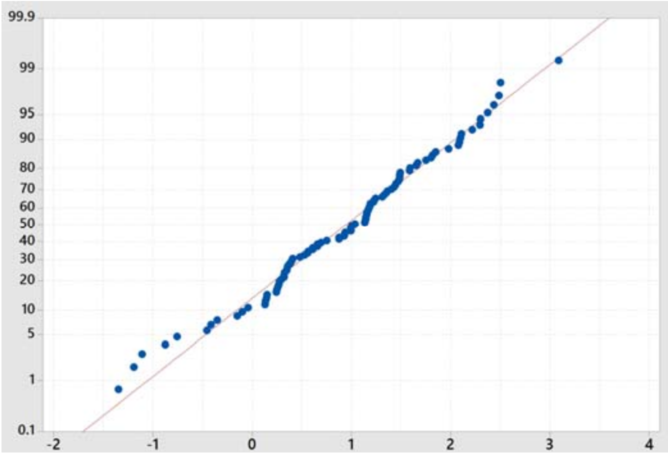
Assume that $\sum_{i=1}^{80} y_i = 74$ and $\sum_{i=81}^{100} y_i = 26$.

(a) Calculate the sample mean for this sample.

$\bar{y}$ =______________

(b) You are told that the sample standard deviation equals 0. Given the information above, is it possible? Explain your answer in a few sentences.

A QQ plot of the observation (Probability Plot in MINITAB) appears as follows:



(c) Explain (in no more than 4 sentences) how a QQ plot is generated and based on the above plot conclude if there is any strong evidence for Normality/Non-Normality of the data.

(d) Assume now that the smallest 10 observations are changed to equal exactly -1 and the largest 10 observations are change to equal exactly +3. Draw a sketch of the revised QQ Plot.

Question 2

A pile of assorted bricks is used to build a wall. The mean height of a brick is 10cm, with a standard deviation of 3cm. Brick heights are assumed to be Normally distributed.

(a) What is the probability that a randomly chosen brick has a height in the range $[7,13]$?

Probability = _______________

(b) A sample of 16 bricks is taken. What is the probability that the sample mean is less than 9cm?

Probability = _______________

(c) If 10 bricks are laid on top of each other, the mean height of the resulting wall is 1m (assuming there are no grout lines between each layer), however the actual height of a wall constructed in this fashion will vary. A finished height of less than 1m is not desirable. What is the minimum number of bricks needed to ensure that the height of the wall will exceed 1m with a probability of at least 99%?

n = _______________

(d) You are now using two bricks with a single layer of grout between them to construct a tiny wall with a desired height of at least 12 cm. Assume grout height is randomly distributed according to a Normal distribution with a mean of 3cm and a standard deviation of s. What is the minimum s, such that the wall height will exceed 12cm with a probability of at least 99%?

s = _______________

Question 3

You are investigating the wear of a bicycle chain using two types of lubricants. You obtain data for the wear using lubricant X on 12 bicycles and the wear using lubricant Y on 10 bicycles. The sample means are 8mm and 5mm respectively and the sample standard deviations are 2.3mm and 2.1mm .

(a) Assume you have reason to believe that the population variances of the wear are the same. Calculate the pooled sample standard deviation.

s_P = _______________

(b) Construct and interpret a 95% confidence interval for the difference in mean wear between the two lubricants.

Confidence Interval=______________________

Interpretation: ____

____

____

(c) You wish to test whether the mean wear is the same for both lubricants. Carry out an hypothesis test for: $H_0 : \mu_X = \mu_Y$ vs. $H_A : \mu_X \neq \mu_Y$, assuming $\alpha = 0.01,$ and state your conclusion.

p-value=____________________

conclusion: ____

____

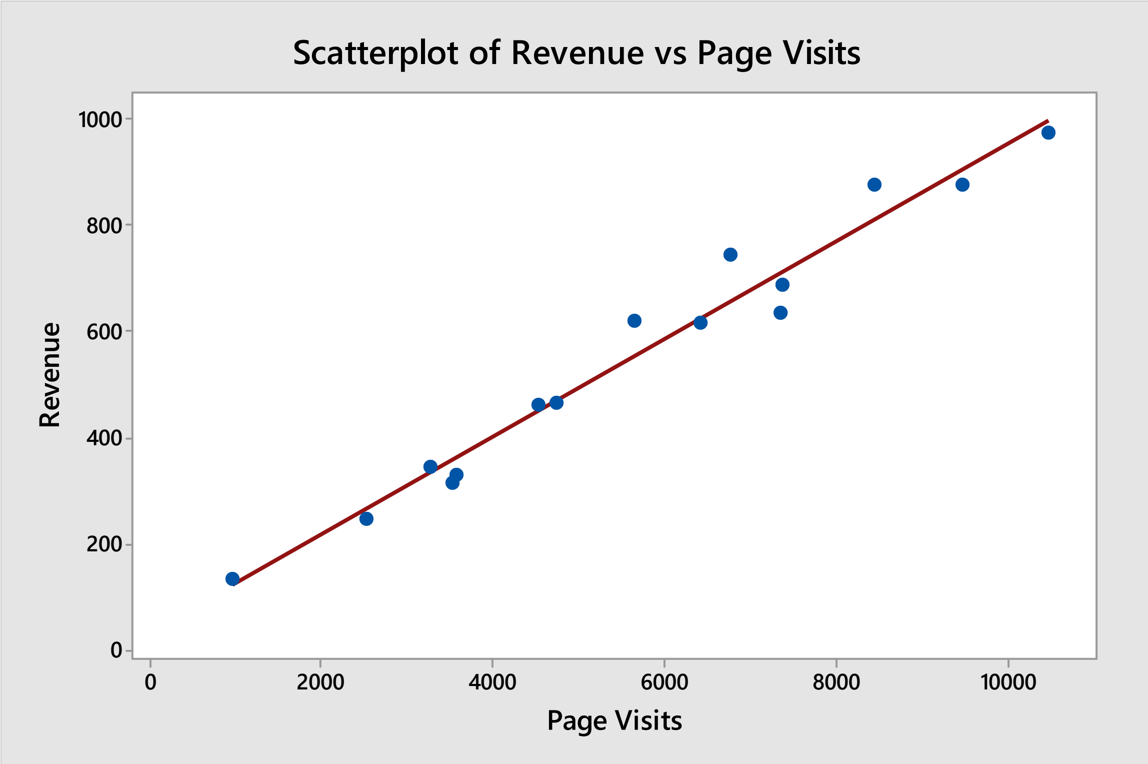
____

(d) Assume that you wish to construct a confidence interval for the wear of lubricant X with a width of no more than 0.5mm and a confidence level of 99%. Assuming the sample standard deviation as above, calculate the minimum sample size needed.

n = _____________

Question 4

A manufacturing company is trying to become e-smart. It has gone on-line and is investigating the weekly revenue obtained from advertising. Over 15 weeks, advertising revenue (in $) is recorded as a function of the number of page visits. The data (together with a least squares line) looks as follows:



(a) Describe the graph above. Does the model,

$$Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j \text{, with } \varepsilon_j \sim N(0, \sigma^2),$$

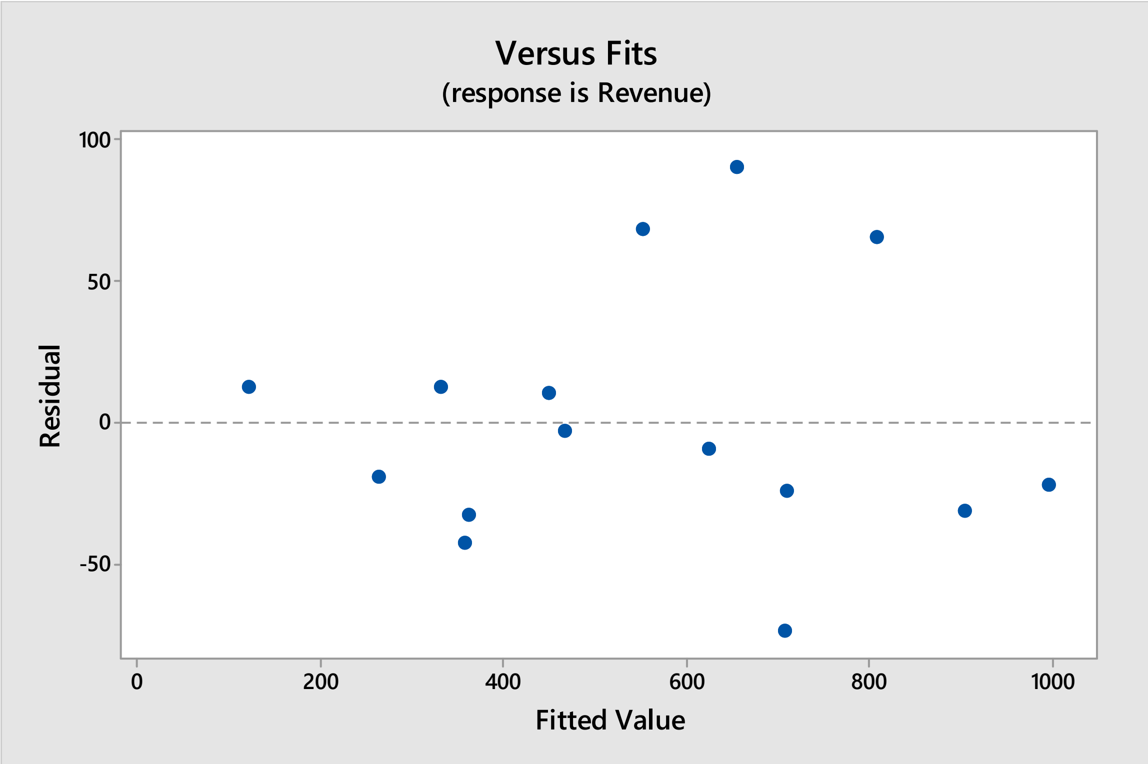
seem reasonable?

(b) In fitting the model above, MINITAB output yields:

```
Coefficients

Term              Coef  SE Coef  T-Value

Constant          32.8     28.9     1.13

Page Visits  0.09208  0.00464    19.84
```

Test the significance of the regression. Namely $H_0 : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$. Use the fact that under the null-hypothesis, the test statistic follows a t-distribution with n-2 degrees of freedom.

(c) In carrying out the hypothesis test above, there are some assumptions made about the distribution of the residuals and their variances.



Based on the above plot, comment on the validity of the corresponding model assumption:

(d) Assume an extra observation is collected with revenue of $832 from a week with only 198 page visits. How would this change the fitted model and output? Comment on the estimates of the slope and the intercept as well on the residual plot.

Formulae and Tables

Summary Statistics:

Suppose $Y_1, Y_2, \cdots, Y_n \sim N(\mu_Y, \sigma_Y^2)$ and $X_1, X_2, \cdots, X_m \sim N(\mu_X, \sigma_X^2)$ are two independent samples. The sample means and sample variances are respectively,

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2,$$

$$\overline{X} = \frac{1}{m}\sum_{i=1}^{m} X_i \qquad S_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(X_i - \overline{X}\right)^2.$$

Note that, $\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = (\sum_{i=1}^{n} Y_i^2) - n\overline{Y}^2$.

The (two-sample) pooled sample variance is $S_P^2 = \frac{(n-1)S_Y^2 + (m-1)S_X^2}{n+m-2}$.

The ECDF function is: $F(t) = \frac{1}{n}\sum_{i=1}^{n} 1\{Y_i \le t\}$.

Sampling Distributions:

$$\frac{\overline{Y} - \mu_Y}{\sigma_Y / \sqrt{n}} \sim Z \qquad \frac{\overline{Y} - \mu_Y}{S_Y / \sqrt{n}} \sim t_{n-1} \qquad \frac{\left(\overline{Y} - \overline{X}\right) - \left(\mu_Y - \mu_X\right)}{S_P\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2} \text{ if } \sigma_X = \sigma_Y .$$

Confidence Intervals for the Mean:

If σ_Y is known, use $\left(\overline{y} \pm z^* \sigma_Y / \sqrt{n}\right)$.

If σ_Y estimated by s_Y ,use $\left(\overline{y} \pm t^* s_Y / \sqrt{n}\right)$.

Confidence Intervals for the Difference in Means for Independent Samples:

If σ_Y and σ_X are known, use $\left(\left(\overline{y} - \overline{x}\right) \pm z^* \sqrt{\frac{\sigma_Y^2}{n} + \frac{\sigma_X^2}{m}}\right)$.

If σ_Y and σ_X are unknown but assumed equal, use $\left(\left(\overline{y} - \overline{x}\right) \pm t^* s_P \sqrt{\frac{1}{n} + \frac{1}{m}}\right)$.

Hypothesis Test Basics:

Type I error: Rejection of the null hypothesis falsely (rejecting the null hypothesis when it is actually true).
Type II error: Non-rejection (retention) of the null hypothesis falsely (not-rejecting the null hypothesis when the alternative hypothesis is true).

$$\Pr(type\,I\,error) = \Pr(reject H_0 \mid H_0 \; holds)$$
$$\Pr(type\,II\,error) = \Pr(not reject H_0 \mid H_A \; holds)$$

Significance level : $\alpha = \Pr(type\;\; I\, error)$
Power = $1 - \Pr(type\;\; II\, error)$

General Hypothesis Test Procedure:

i. Write down the null and alternative hypotheses.
ii. Select an appropriate test statistic for the test and compute it based on data.
iii. Sketch the distribution of the test statistic and mark the observed value on the plot (also the “opposite” value if the test is two-sided).
iv. Compute the tail area (or bounds for the tail area), which gives the p-value (multiply the area by 2 for a two-sided alternative hypothesis).
v. State the conclusion: If computationally possible, report the p-value; otherwise, compare the test statistic with the critical value.

Hypothesis Tests for the Mean:

To test against $H_0 : \mu = \mu_0$: If σ_Y is known, use: $\dfrac{(\bar{y} - \mu_0)}{\sigma_Y / \sqrt{n}}$ and the standard Normal distribution; otherwise, use: $\dfrac{(\bar{y} - \mu_0)}{s_Y / \sqrt{n}}$ and the t_{n-1} – distribution.

Hypothesis Tests for the Difference in Means for Independent Samples:

To test against $H_0 : \mu_Y = \mu_X$: If σ_Y and σ_X are known, use $\dfrac{\bar{y} - \bar{x}}{\sqrt{\dfrac{\sigma_Y^2}{n} + \dfrac{\sigma_X^2}{m}}}$ as a test statistic.

If σ_Y and σ_X are unknown but assumed equal, use $\dfrac{\bar{y} - \bar{x}}{s_P \sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$ as a test statistic.

Single Factorial Models (analysed through ANOVA):

For factor levels i=1,…,k, $Y_{i,j} = \mu_i + \varepsilon_{i,j}$, with $\varepsilon_{i,j} \sim N(0, \sigma^2)$.

Simple Linear Regression (estimated through least squares):

$$Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j \text{, with } \varepsilon_j \sim N(0, \sigma^2).$$

Standard Normal Cumulative 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	.2297	.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

This table was generated using the "CDF" command in Minitab.

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	.9773	.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	.9983	.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	.9996	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

t-Distribution Quantiles

ν	$Q(.9)$	$Q(.95)$	$Q(.975)$	$Q(.99)$	$Q(.995)$	$Q(.999)$	$Q(.9995)$
1	3.078	6.314	12.706	31.821	63.657	318.317	636.607
2	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.849
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

This table was generated using the "INVCDF" command in Minitab.

END OF EXAMINATION