Solutions to Sample Exam 1

Problem 1-A

We know that for any normal model, 68% of values are within 1 standard deviation of the mean, and 90% of values are within 1.645 standard deviations of the mean. We assume the mean is the midpoint of the two intervals which is 30. Based on the first interval the standard deviation would therefore be 10. If we check the second interval, to determine what the 1.645 standard deviation range is we get:

$$[30 - 1.645(10), \quad 30 + 1.645(10)]$$

$$[13.55, \quad 46.45]$$

This interval closely matches the second interval of [13.5, 46.5]. Therefore our assumption that the mean is 30 and the standard deviation is 10 is correct. We know that 95% of values will lie within 1.96 standard deviations of the mean, which will be in the following range:

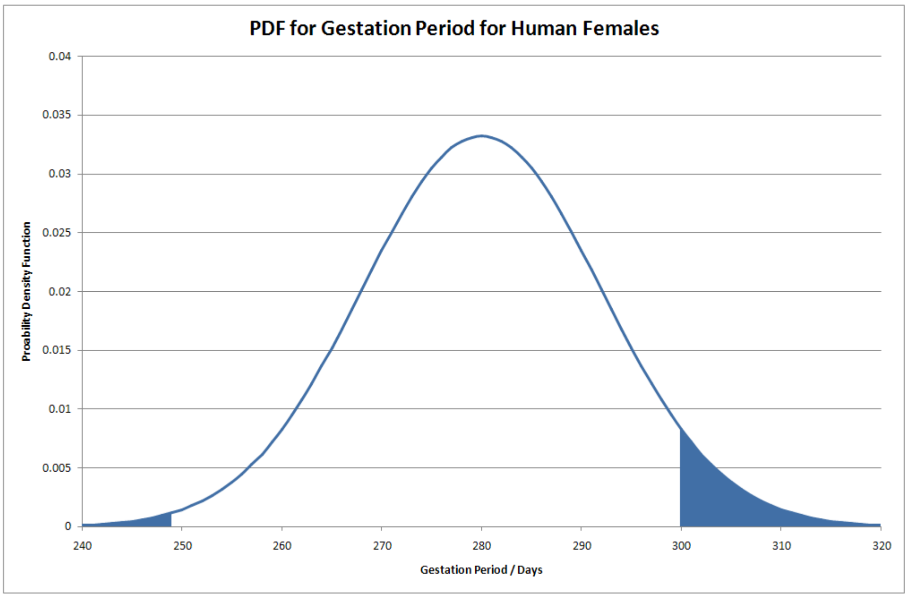
$$[30 - 1.96(10), \quad 30 + 1.96(10)]$$

$$[10.4, \quad 49.6]$$

Therefore the answer is **D**; 95% of the values in the population will be between 10.4 and 49.6

Problem 1-B

We want to determine the probability that the baby was conceived when the alleged father was present. This corresponds to the area under the curve as shown below.



The limits are 249 and 300. However we must standardise these limits with respect to the mean and standard deviation of the gestation period. We find the standardised limits using the following equation:

$$Z = \frac{y - \mu}{\sigma}$$

Z	=	Standard Normal Variable
y	=	Lower Limit (249) or Upper Limit (300)
μ	=	Mean value for gestation period (280)
σ	=	Standard Deviation for gestation period (12)

The standardised lower and upper limits will therefore be given by:

$$SLL = \frac{249 - 280}{12} \qquad SUL = \frac{300 - 280}{12}$$

$$SLL = -2.58 \qquad SUL = 1.67$$

The area under the curve is given by the below equation where Z(2.58) and Z(1.67) are the values obtained from the Z-Table(given on the second last page) for 2.58 and 1.67 respectively.

$$Area = [1 - Z(2.58)] + [1 - Z(1.67)]$$

$$Area = [1 - 0.9951] + [1 - 0.9525]$$

$$Area = 2 - 0.9951 - 0.9525$$

$$Area = 0.0524$$

Therefore the chances that the child was conceived when the father was present is only 5.24%. This means that there is approximately a 95% chance that the alleged father is in fact not the father.

Problem 1-C

- From the AIAG chart on the following page we obtain the following for a subsample size of 4:
  - $A_2 = 0.73$
  - $D_3 = 0$
  - $D_4 = 2.28$

The control limits for the $\bar{X}$ control chart are given by the following equations:

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} \qquad UCL_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$$

$\bar{\bar{X}}$ = mean
$\bar{R}$ = average range

We are told the mean is 140.17 and the average range is 1.11, and since we know the value of A_2 is 0.73 for a subsample size of 4, the controls limits for the $\bar{X}$ control chart will be:

$$LCL_{\bar{X}} = 140.17 - (0.73)(1.11) \qquad UCL_{\bar{X}} = 140.17 + (0.73)(1.11)$$

$$LCL_{\bar{X}} = 140.17 - (0.73)(1.11) \qquad UCL_{\bar{X}} = 140.17 + (0.73)(1.11)$$

$$LCL_{\bar{X}} = 140.17 - 0.8103 \qquad UCL_{\bar{X}} = 140.17 + 0.8103$$

$$LCL_{\bar{X}} = 139.3597 \qquad UCL_{\bar{X}} = 140.9803$$

The control limits for the R control chart are given by the following equations:

$$LCL_R = D_3\bar{R} \qquad UCL_R = D_4\bar{R}$$

We are told the range is 1.11, and since we know the values of D_3 and D_4 are 0 and 2.28 respectively, for a sample size of 4, the controls limits for the R control chart will be:

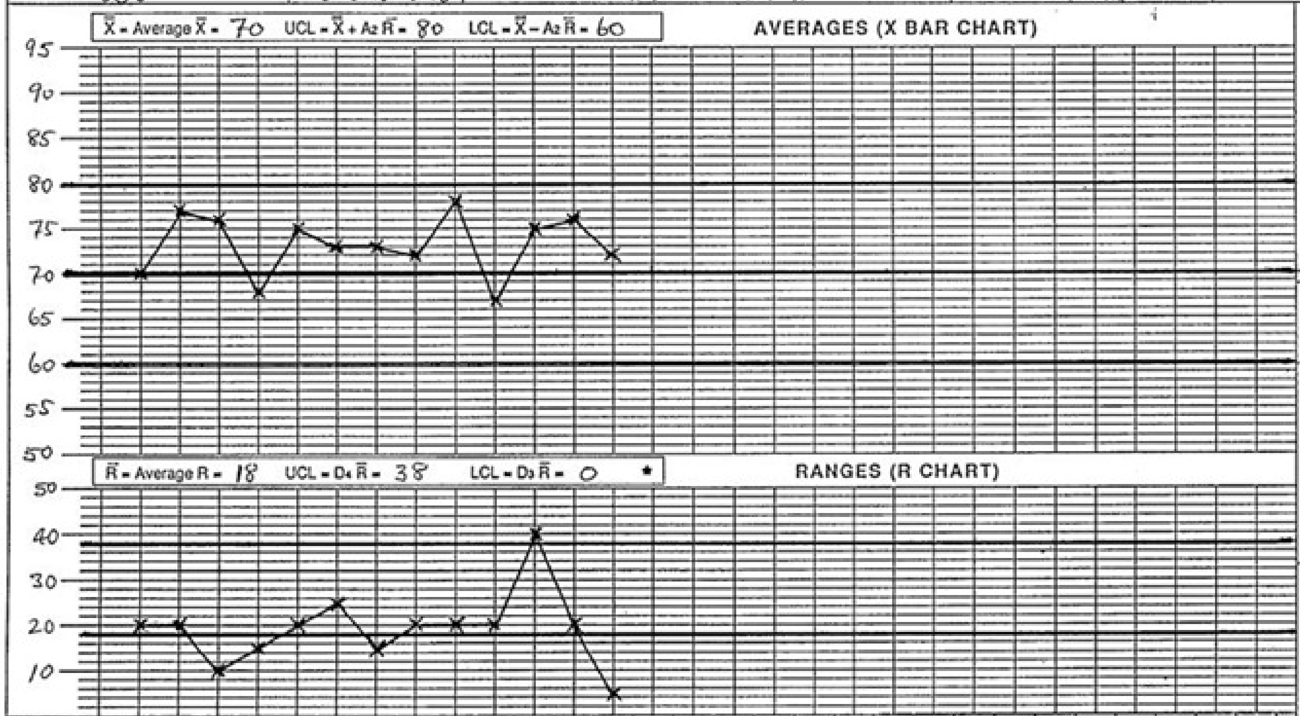
$$LCL_R = (0)(1.11) \qquad UCL_R = (2.28)(1.11)$$

$$LCL_R = 0 \qquad UCL_R = 2.5308$$

Also shown on the charts in blue are the central lines. The central line for the $\bar{X}$ control chart is simply $\bar{\bar{X}}$ and the central line for the R control chart is $\bar{R}$.

CONTROL CHART

PLANT: CHICAGO	DEPT: .105	OPERATION: BEND CLIP	DATE CONTROL LIMITS CALCULATED: 5-4	ENGINEERING SPECIFICATION: 50 to 90 mm	PART NO.: E2B3 123
MACH. NO.: 030	DATES: 6-8 to 6-16	CHARACTERISTIC: CLIP GAP DIM "A"		SAMPLE SIZE/FREQUENCY: 5 / 2 HRS	PART NAME: RETAINER



DATE / TIME		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
READINGS 1		65	75	75	60	70	60	75	60	65	60	80	85	70	65	90
2		70	85	80	70	75	75	80	70	80	70	75	75	70	70	80
3		65	75	80	70	65	75	65	80	85	60	90	85	75	85	80
4		65	85	70	75	85	85	75	75	85	80	50	65	75	75	75
5		85	65	75	65	80	70	70	75	75	65	80	70	70	60	85
SUM		350	385	380	340	375	365	365	360	390	335	375	380	360		
$\bar{X}$ = SUM / NO. OF READINGS		70	77	76	68	75	73	73	72	78	67	75	76	72		
R = HIGHEST − LOWEST		20	20	10	15	20	25	15	20	20	20	40	20	5		

ACTION
ON SPECIAL CAUSES

- ANY POINT OUTSIDE OF THE CONTROL LIMITS
- A RUN OF 7 POINTS ALL ABOVE OR ALL BELOW THE CENTRAL LINE
- A RUN OF 7 POINTS UP OR DOWN
- ANY OTHER OBVIOUSLY NON-RANDOM PATTERN

ACTION INSTRUCTIONS

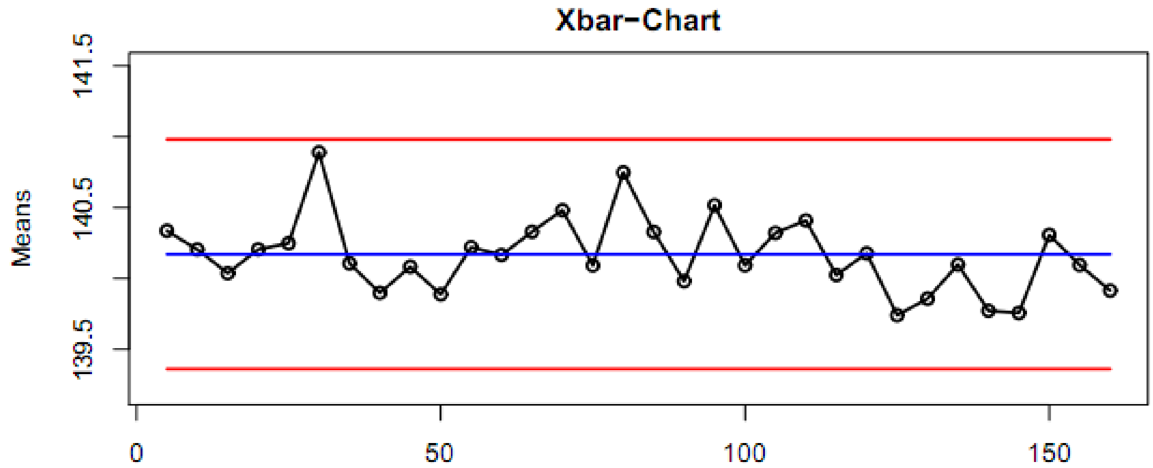
1.
2.
3.
4.
5.

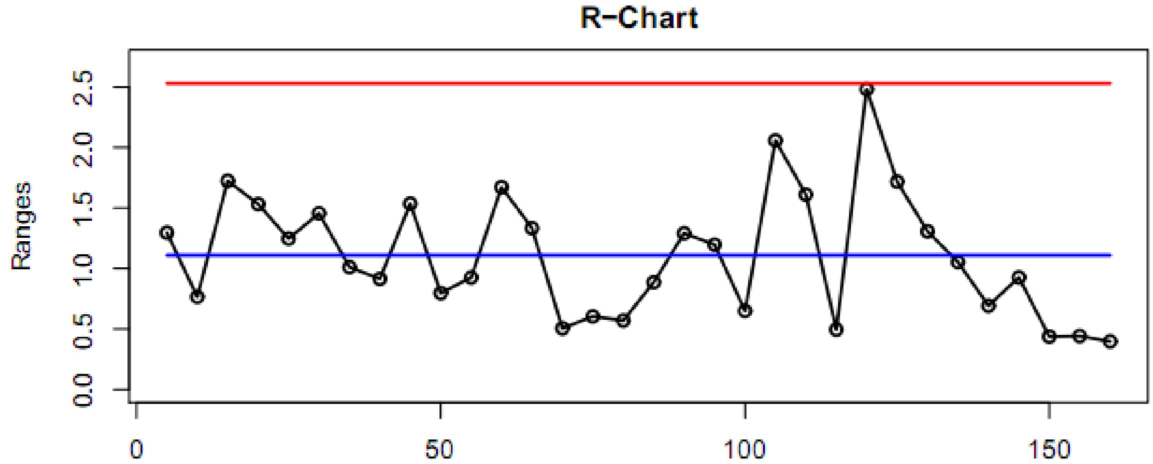
SUBGROUP SIZE	A_2	D_3	D_4
2	1.88	*	3.27
3	1.02	*	2.57
4	.73	*	2.28
5	.58	*	2.11
6	.48	*	2.00
7	.42	.08	1.92
8	.37	.14	1.86
9	.34	.18	1.82
10	.31	.22	1.78

THE PROCESS MUST BE IN CONTROL BEFORE CAPABILITY CAN BE DETERMINED.

* For sample sizes of less than seven, there is no lower control limit for ranges.

The control limits are draw below for the $\bar{X}$ control chart and the Range control chart. We don't include the lower limit for the Range control chart, since we will never obtain a negative value for the range.





- The mean value for the brick widths is equivalent to $\bar{X}$ which is 140.17. To determine the standard deviation of the process we use the following equation where n is the subsample size, in this case 4:

$$\hat{\sigma} = \frac{A_2\sqrt{n}}{3}\bar{R}$$

$$\hat{\sigma} = \frac{0.73\sqrt{4}}{3}1.11$$

$$\hat{\sigma} = 0.5402$$

Therefore the three-sigma band is given by:

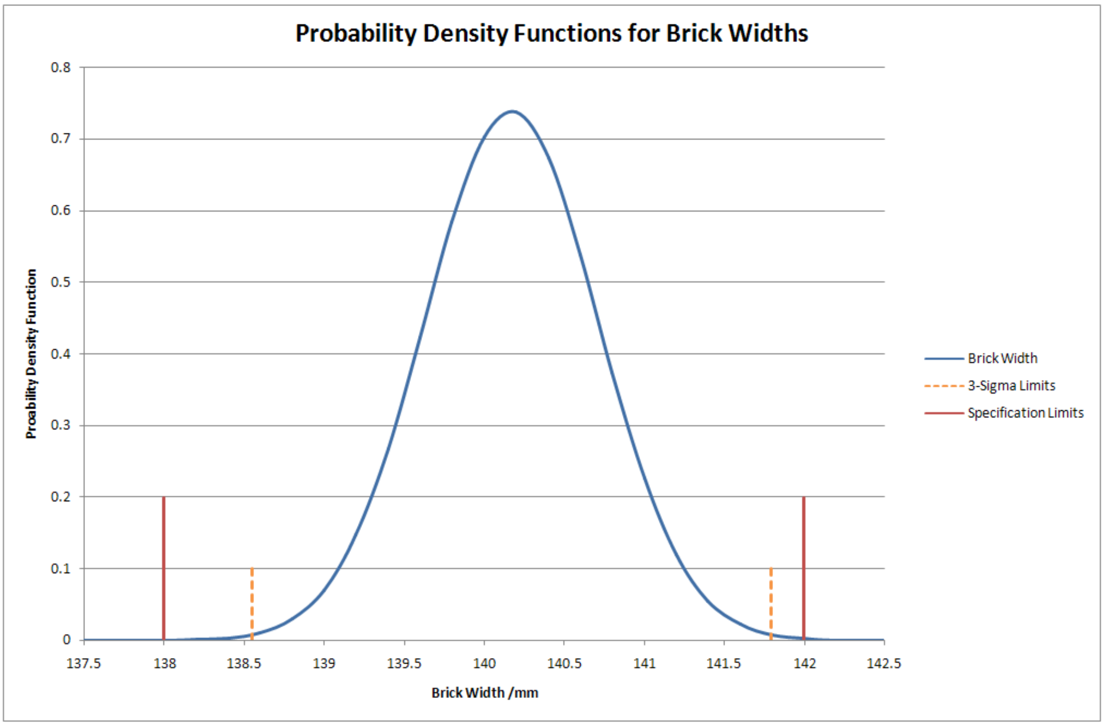
$$[\mu - 3\hat{\sigma}, \mu + 3\hat{\sigma}]$$

$$[140.17 - 3(0.5402),\ \ 140.17 + 3(0.5402)]$$

$$[140.17 - 1.6206,\ \ 140.17 + 1.6206]$$

$$[138.5494,\ \ 141.7906]$$

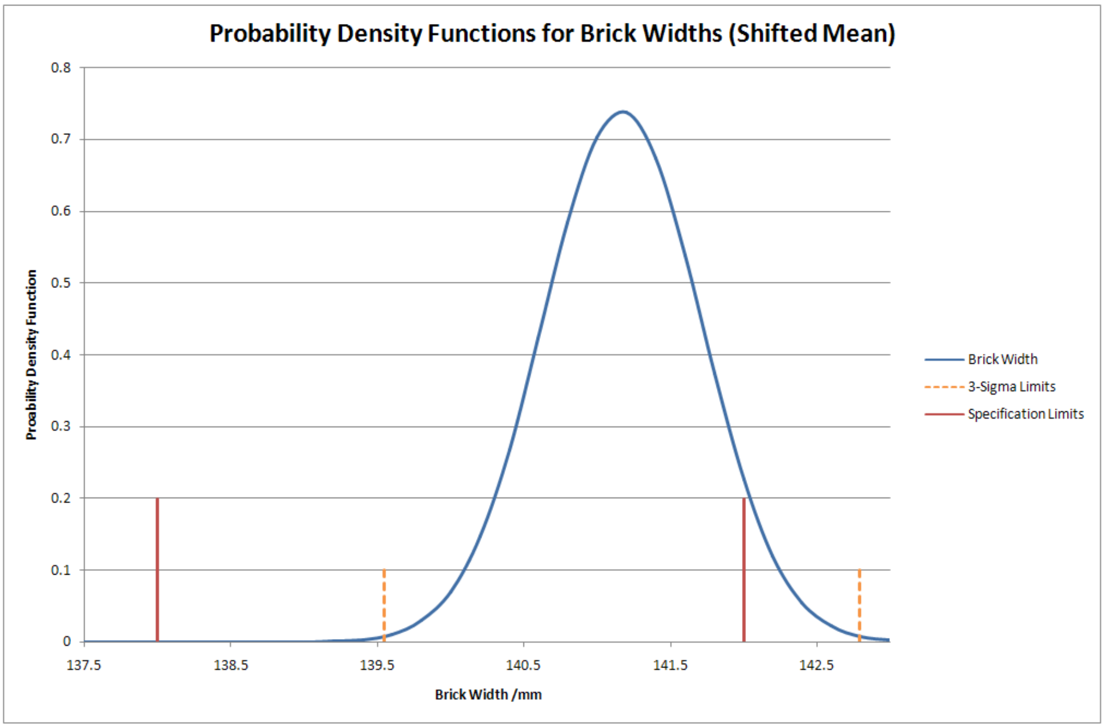
We see that the target width for these bricks is 140mm ± 2mm i.e. within the interval of [138, 142]. Because the three-sigma band falls inside the target interval, the target specifications are met i.e. we will have a non-conformance rate of less than 3 in 1,000.



- The new mean would be 141.17 and the three-sigma band would shift to the right by 1 to:

  $[139.5494, \quad 142.7906]$

  The target interval for the bricks is [138,142]. Because the three-sigma band does not fall inside the target interval, the target specifications are not met i.e. we will have a non-conformance rate of greater than 3 in 1,000. The upper specification limit is well below the upper three-sigma limit.



- The new mean is 141.17 and the standard deviation has doubled to 1.0804. Therefore the three-sigma band is given by:

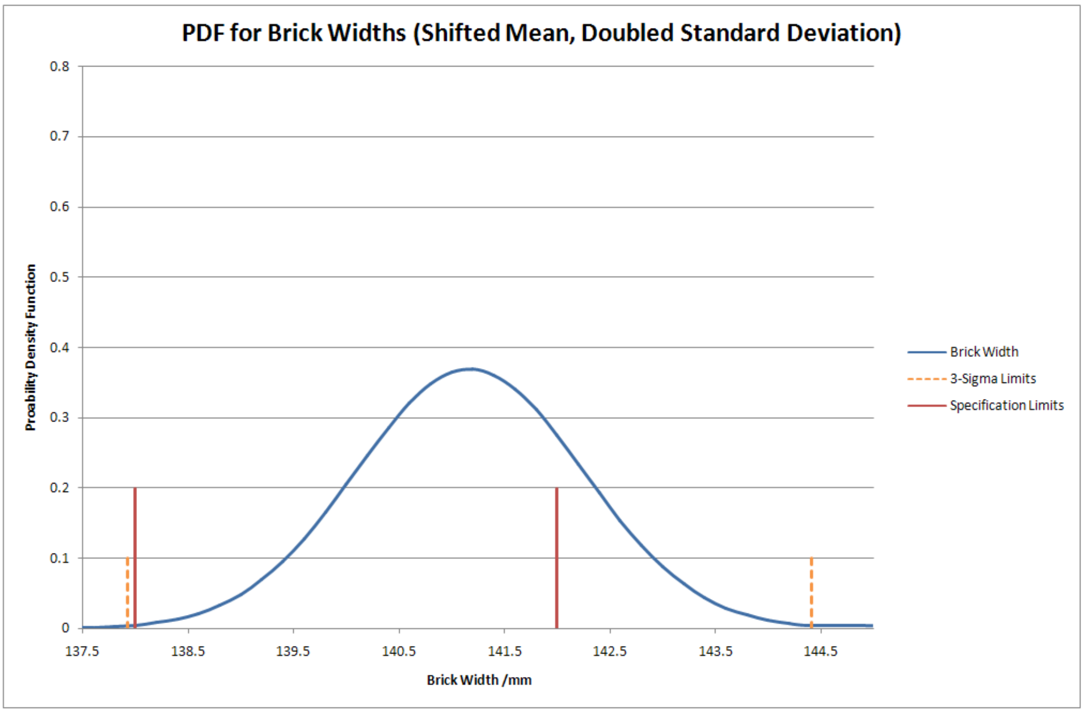
  $[\mu - 3\hat{\sigma}, \mu + 3\hat{\sigma}]$

  $[141.17 - 3(1.0804),\ \ 141.17 + 3(1.0804)]$

  $[141.17 - 3.2412,\ \ 141.17 + 3.2412]$

  $[137.9288,\ \ 144.4112]$

  The target interval for the bricks is [138,142]. Because the three-sigma band does not fall within the target interval, the target specifications are not met i.e. we will have a non-conformance rate of greater than 3 in 1,000. The upper specification limit is well below the upper three-sigma limit. The lower specification limit is just above the lower three-sigma limit.



Problem 2-A

- We are checking the null hypothesis that the means of the two samples are equal i.e.

  H_O : $\mu_1 = \mu_2$

  H_a : $\mu_1 \neq \mu_2$

  Since the null hypothesis we are checking only uses the equality symbol (=) the test is two-sided.

- We will use a 5% significance level

- We use the following test statistic:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\bar{X}_1$	=	Sample mean of first dataset (660.67)
$\bar{X}_2$	=	Sample mean of second dataset (501.83)
n_1	=	Sample size of first dataset (12)
n_2	=	Sample size of second dataset (6)
S_p	=	Pooled standard deviation
T	=	Test statistic

The pooled standard deviation is given by:

$$S_p = \sqrt{\frac{{s_1}^2(n_1 - 1) + {s_2}^2(n_2 - 1)}{(n_1 - 1) + (n_2 - 1)}}$$

s_1	=	Standard deviation of first dataset (111.47)
s_2	=	Standard deviation of second dataset (131.99)

- Because the sample size is less than 30, this statistic follows the Student's t-distribution with ($n_1 + n_2 - 2$) = 16 degrees-of-freedom. We obtain the critical value for this test from the Student's t-distribution table. We know that the significance level is 5%, and since the test is two-sided when we consult the table we check for just one tail which will be 2.5%. Therefore the critical value for this test from the Student's t-distribution table is 2.120.

- To evaluate the test statistic we first need to determine the pooled standard deviation:

$$S_p = \sqrt{\frac{(111.47)^2(12-1) + (131.99)^2(6-1)}{12+6-2}}$$

$$S_p = \sqrt{\frac{11(111.47)^2 + 5(131.99)^2}{16}}$$

$$S_p = \sqrt{\frac{11(12425.5609) + 5(17421.3601)}{16}}$$

$$S_p = \sqrt{\frac{223787.9704}{16}}$$

$$S_p \approx 118.27$$

Now we work out the test statistic:

$$T = \frac{660.67 - 501.83}{118.27\sqrt{\frac{1}{12} + \frac{1}{6}}}$$

$$T = \frac{158.84}{118.27\sqrt{\frac{1}{12} + \frac{2}{12}}}$$

$$T = \frac{158.84}{118.27\sqrt{\frac{3}{12}}}$$

$$T = \frac{158.84}{118.27}\sqrt{\frac{12}{3}}$$

$$T = \frac{158.84}{118.27}\sqrt{4}$$

$$T = 2\left(\frac{158.84}{118.27}\right)$$

$$T \approx 2.69$$

- Out test statistic is 2.69. In order to reject the null hypothesis the value should be either below the negative critical value of -2.120 or above the positive critical value of 2.120. Because the test statistic is outside the range of [-2.210, 2.210] we reject the null hypothesis.

- To get the *p*-value we need to determine the area under the Student's t-distribution curve with 16 degrees-of-freedom, outside the range of [-2.69, 2.69]. When we check the table we see that a value of 2.583 corresponds to a single tail area of 1% and 2.921 corresponds to a single tail area of 0.5%. Since the value of 2.69 lies roughly midway between these two values, 2.69 should correspond to a single tail area of 0.75%. We want the area of two tails, so the *p*-value will be approximately 1.5%. We expected the *p*-value to be less than the significance level of 5% since we already decided to reject the null hypothesis because the test statistic of 2.69 lies outside the range of [-2.210, 2.210].

Problem 2-B

1. The 95% confidence interval will be given by the following equation:

$$\left[\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$\hat{p}$ = Observed error proportion (30/150 = 0.2)
n = Sample size (150)

$$\left[0.2 - 1.96\sqrt{\frac{0.2(1-0.2)}{150}}, \quad 0.2 + 1.96\sqrt{\frac{0.2(1-0.2)}{150}}\right]$$

$$\left[0.2 - 1.96\sqrt{\frac{0.2(0.8)}{150}}, \quad 0.2 + 1.96\sqrt{\frac{0.2(0.8)}{150}}\right]$$

$$\left[0.2 - 1.96\sqrt{\frac{0.16}{150}}, \quad 0.2 + 1.96\sqrt{\frac{0.16}{150}}\right]$$

$$[0.2 - 1.96(0.0327), \quad 0.2 + 1.96(0.0327)]$$

$$[0.2 - 0.064, \quad 0.2 + 0.064]$$

$$[0.136, \quad 0.264]$$

Based on the sample data we are 95% confident that the actual transaction error rate lies between 13.6% and 26.4%

2. The 95% confidence interval will be given by the following equation:

$$\left[\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$\hat{p}$ = Observed error proportion (120/600 = 0.2)
n = Sample size (600)

$$\left[0.2 - 1.96\sqrt{\frac{0.2(1-0.2)}{600}}, \quad 0.2 + 1.96\sqrt{\frac{0.2(1-0.2)}{600}}\right]$$

$$\left[0.2 - 1.96\sqrt{\frac{0.2(0.8)}{600}}, \quad 0.2 + 1.96\sqrt{\frac{0.2(0.8)}{600}}\right]$$

$$\left[0.2 - 1.96\sqrt{\frac{0.16}{600}}, \quad 0.2 + 1.96\sqrt{\frac{0.16}{600}}\right]$$

$$[0.2 - 1.96(0.0163), \quad 0.2 + 1.96(0.0163)]$$

$$[0.2 - 0.031948, \quad 0.2 + 0.031948]$$

$$[0.168, \quad 0.232]$$

Based on the larger sample data we are 95% confident that the actual transaction error rate lies between 16.8% and 23.2%. The confidence interval width is exactly half that in the first part of this question; multiplying the sample size by a factor of 4 means that we half the confidence interval width.

3. The 95% confidence interval will be given by the following equation:

$$\left[\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$\hat{p}$	=	Observed error proportion (30/600 = 0.05)
n	=	Sample size (600)

$$\left[0.05 - 1.96\sqrt{\frac{0.05(1-0.05)}{600}}, \quad 0.05 + 1.96\sqrt{\frac{0.05(1-0.05)}{600}}\right]$$

$$\left[0.05 - 1.96\sqrt{\frac{0.05(0.95)}{600}}, \quad 0.05 + 1.96\sqrt{\frac{0.05(0.95)}{600}}\right]$$

$$\left[0.05 - 1.96\sqrt{\frac{0.0475}{600}}, \quad 0.05 + 1.96\sqrt{\frac{0.0475}{600}}\right]$$

$$[0.05 - 1.96(0.0089), \quad 0.05 + 1.96(0.0089)]$$

$$[0.05 - 0.017444, \quad 0.05 + 0.017444]$$

$$[0.033, \quad 0.067]$$

Based on the larger sample data, with a reduced proportion of error transactions, we are 95% confident that the actual transaction error rate lies between 3.3% and 6.7%.

4. If we increase the sample size by a factor of *m,* the confidence interval width will decrease by a factor of $\sqrt{m}$. In the second part of this question we increased the sample size by a factor of 4, and the confidence interval width was halved.

Problem 3-A

B – The accuracy of predicting the number of homepage visitors when the information about the promotion spending is not available is reduced by a factor of:

$$\sqrt{\frac{(1-r^2)(n-1)}{n-2}}$$

This is approximately:

$$\sqrt{1-r^2}$$

We are told in the problem that the correlation coefficient *r* is 0.84. Therefore the reduction factor is:

$$\sqrt{1-(0.84)^2}$$

$$\sqrt{1-0.7056}$$

$$\sqrt{0.2944}$$

$$0.54$$

Therefore the correct answer is B. Also, it is worth noting that with regard to choice C, it is the R-squared coefficient that tells us the percentage variation of the dependence of one variable on another.

Problem 3-B

- The response variable is the number of computers produced on a weekly basis. The explanatory variable is the associated weekly production cost.

- The simple regression model will have the form:

  $$Y = a + bX + \epsilon$$

a	=	Intercept (2272.07)
b	=	Slope (51.6612)
ϵ	=	Standard deviation of error term (198.6)

- We know the equation for the line will be given by:

  $$Y = 2272.07 + 51.6612X$$

  We need to pick two points, firstly *X = 22*

  $$X = 22$$

  $$Y = 2272.07 + 51.6612(22)$$

  $$Y \approx 3409$$

  Now pick a second point, *X = 45*

  $$X = 45$$

  $$Y = 2272.07 + 51.6612(45)$$

  $$Y \approx 4597$$

  We now have two points [22, 3409] and [45, 4597] so we can plot the line. The points and the fitted line are shown on the following page.

- The actual cost associated with week 12 is 4,158. The predicted cost using the simple linear regression model will be given by the below equation where X is the number of computers produced in week 12 which is 41:

  $$Y_{W12} = 2272.07 + 51.6612X$$

  $$Y_{W12} = 2272.07 + 51.6612(41)$$

  $$Y_{W12} = 2272.07 + 51.6612(41)$$

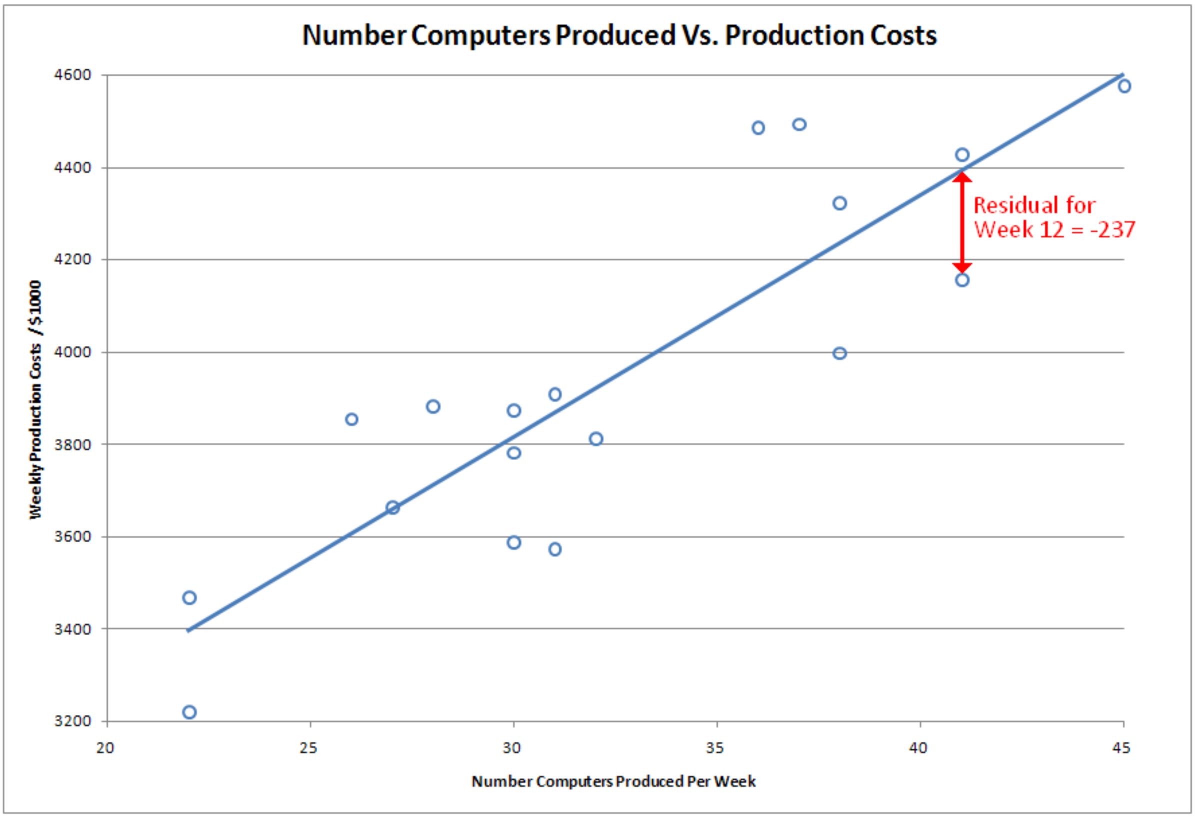
  $$Y_{W12} \approx 4390$$

The residual is the difference between the actual cost and the predicted cost:

$$\epsilon_{W12} = 4158 - 4390$$

$$\epsilon_{W12} = -237$$

This residual is shown on the below graph.



- The standardised test statistic to determine if the slope is non-zero is given by:

$$\frac{\hat{\beta}}{SE(\hat{\beta})}$$

$\hat{\beta}$	=	Estimation of Slope (51.6612)
$SE(\hat{\beta})$	=	Standard Error of Estimation of Slope (7.347)

For our data the slope is therefore:

$$\frac{51.6612}{7.347}$$

$$7.03$$

The standardised test statistic to determine if the intercept is non-zero is given by:

$$\frac{\hat{\alpha}}{SE(\hat{\alpha})}$$

$\hat{\alpha}$ = Estimation of Intercept (2272.07)
$SE(\hat{\alpha})$ = Standard Error of Estimation of Intercept (243.3)

For our data the slope is therefore:

$$\frac{2272.07}{243.3}$$

$$9.34$$

The critical value for both tests at a 5% significance level corresponds to the area under the Student's t-distribution curve with *n – 2 = 16* degrees-of-freedom for an area of 2.5%. This gives us a critical value of 2.12. Since both test statistics (7.03, 9.34) are greater than the critical value of 2.12 we reject the hypotheses that the slope is zero and the intercept is zero.

- If we were to produce 40 computers, then we use the simple regression model with X = 40

$$Y = 2272.07 + 51.6612X$$

$$Y = 2272.07 + 51.6612(40)$$

$$Y \approx 4339$$

We know the standard deviation is 198.6, therefore when we add two-sigma limits:

$$[4339 \pm 2(198.6)]$$

$$[4339 \pm 397.2]$$

$$[3942, 4736]$$

Therefore if we plan to manufacture 40 computers next week we will be 95% (corresponds to two-sigma limits) confident that the cost will be between \$3,942,000 and \$4,736,000.

- Since the value of 5,000 is outside the two-sigma band, we must conclude that this value indicates some special circumstances that cannot be explained by chance variation in the linear regression model. The case must be examined for other potentially controllable causes of variation.

Problem 4

1. From the computer readouts we obtain the following data:

    - Regression Model 1 (Acetic) $\hat{Y} = -61.499 + 15.648X_a$
    - Regression Model 2 (Hydrogen Sulphide) $\hat{Y} = -9.7688 + 5.7761X_h$
    - Regression Model 3 (Lactic) $\hat{Y} = -29.859 + 37.720X_l$

2. The H_2S variable explains most of the variability of the taste variable. We arrive at this conclusion because the R-squared coefficient for this variable is the largest at 57.12%.

3. All three regression models are statistically significant because the *p*-values for their slopes are smaller than 5% (0.00166, 0.00000137, 0.0000141). This means we must reject the null hypothesis that the slopes are zero. Alternatively we could compare the test statistics (3.481, 6.107, 5.249) and compare these values with the critical value from the Student's t-distribution table for 28 degrees-of-freedom[1] for a single tail of 2.5% which gives a value of 2.048. The test statistics are all greater than the critical value of 2.048, meaning we reject the null hypothesis that the slopes are zero.

4. The multivariate regression model will be of the form:

    $$\hat{Y} = -\hat{a} + \hat{b}_1X_1 + \hat{b}_2X_2 + \cdots + \hat{b}_dX_d$$

$\hat{a}$	=	Intercept
X_i	=	Explanatory Variable
d	=	Index for total number of explanatory variables
$\hat{Y}$	=	Response Variable

    In model 1 we have three explanatory variables (Acetic, Hydrogen Sulphide, Lactic) and a single response variable (Taste). Looking at the data from the computer readout, the fitted multivariate linear regression for model 1 is therefore:

    $$\hat{Y} = -28.8768 + 0.3277X_a + 3.9118X_h + 19.6705X_l$$

    In model 2 we have two explanatory variables (Hydrogen Sulphide, Lactic) and a single response variable (Taste). Looking at the data from the computer readout, the fitted multivariate linear regression for model 2 is therefore:

    $$\hat{Y} = -27.592 + 3.946X_h + 19.887X_l$$

[1] Since our sample size is 30 and we are comparing 2 variables (one explanatory, and the response variable) are degrees-of-freedom will be given by 30 – 2 = 28.

5. Looking at the computer readout, we check the *p*-values for each explanatory variable for each of the two models. If a *p*-value is less than 5% it means that the variable is significant in explaining the taste variable. Therefore, for the first model Hydrogen Sulphide and Lactic are significant in explaining the taste variable, while Acetic is not. For the second model both explanatory variables are significant in explaining the response variable. The results are presented below in table format.

Model	Explanatory Variable	*p*-value	Significance Level	Significant
1	Acetic	94.198%	5%	no
1	Hydrogen Sulphide	0.425%	5%	yes
1	Lactic	3.108%	5%	yes
2	Hydrogen Sulphide	0.174%	5%	yes
2	Lactic	1.885%	5%	yes

Alternatively, we could compare the test statistics (t value from the computer readout) and see if this number lies outside the range defined by:

$$\left[-t_{\frac{\alpha}{2},n-2}, t_{\frac{\alpha}{2},n-2}\right]$$

$t_{\frac{\alpha}{2},n-1}$ = Value from Student's t-distribution for area of α/2 and n-2 degrees-of-freedom

α = Significance Level (5%)

n = Sample Size (30)

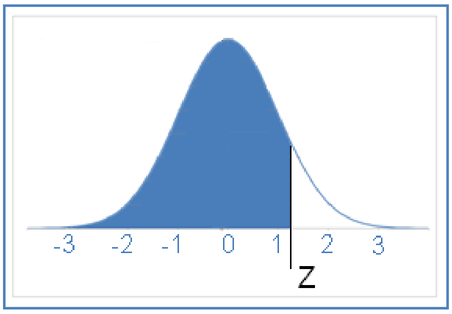
The value from the Student's t-distribution for an area of 2.5% and 28 degrees-of-freedom is 2.048. If the test statistic lies outside the range [-2.048, 2.048] then the associated explanatory variable is significant in explaining the response variable. We notice that only Acetic from model 1 is inside the range [-2.048, 2.048] and hence isn't significant in explaining the response variable.

6. The R-squared value well inform us how much of the variability of the taste variable is explained by the model. For the first multivariate linear regression model the R-squared value is 65.18% and for the second, it is 65.17%. Therefore model 1 is slightly better at explaining the variability of the response variable. Both of these multivariate models explain more of the variability in the response variable, than any of the individual simple linear regression models. The results for all the models are shown below in table form:

Model	R-squared value
Simple Linear Regression Model 1	30.2%
Simple Linear Regression Model 2	57.12%
Simple Linear Regression Model 3	49.59%
Multivariate Linear Regression Model 1	65.18%
Multivariate Linear Regression Model 2	65.17%

7. We can assume the R-squared values from the two multivariate models to be equal, so we would pick the second multivariate model since it involves only two explanatory variables. The Acetic variable, included in the first model, has been shown to be redundant i.e. it is not significant in explaining the response variable.

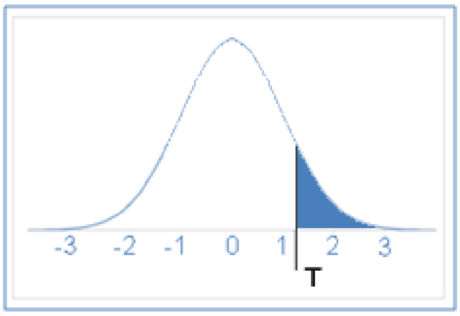
8. No. We do not observe any exceptionally large or small values (residuals), thus there is no evidence of a serious problem with the fit.



Z-Table (Area Under the Curve[2])

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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

[2] Values used in the problems are boxed in red.



Student’s t-Distribution Table[3]

Area	10%	5%	4%	2.5%	2.0%	1%	0.5%	0.25%	0.1%	0.05%
DoF										
1	3.078	6.314	7.916	12.706	15.894	31.821	63.656	127.321	318.289	636.578
2	1.886	2.920	3.320	4.303	4.849	6.965	9.925	14.089	22.328	31.600
3	1.638	2.353	2.605	3.182	3.482	4.541	5.841	7.453	10.214	12.924
4	1.533	2.132	2.333	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.191	2.571	2.757	3.365	4.032	4.773	5.894	6.869
6	1.440	1.943	2.104	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.046	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.004	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	1.973	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	1.372	1.812	1.948	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	1.363	1.796	1.928	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	1.356	1.782	1.912	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	1.350	1.771	1.899	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	1.345	1.761	1.887	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	1.341	1.753	1.878	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	1.337	1.746	1.869	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	1.333	1.740	1.862	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	1.330	1.734	1.855	2.101	2.214	2.552	2.878	3.197	3.610	3.922
19	1.328	1.729	1.850	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	1.325	1.725	1.844	2.086	2.197	2.528	2.845	3.153	3.552	3.85
21	1.323	1.721	1.840	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	1.321	1.717	1.835	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	1.319	1.714	1.832	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	1.318	1.711	1.828	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	1.316	1.708	1.825	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	1.315	1.706	1.822	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	1.314	1.703	1.819	2.052	2.158	2.473	2.771	3.057	3.421	3.689
28	1.313	1.701	1.817	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	1.311	1.699	1.814	2.045	2.150	2.462	2.756	3.038	3.396	3.660
30	1.310	1.697	1.812	2.042	2.147	2.457	2.75	3.030	3.385	3.646
inf	1.282	1.654	1.751	1.960	2.054	2.326	2.576	2.807	3.090	3.290

[3] Values used in the problems are boxed in red.