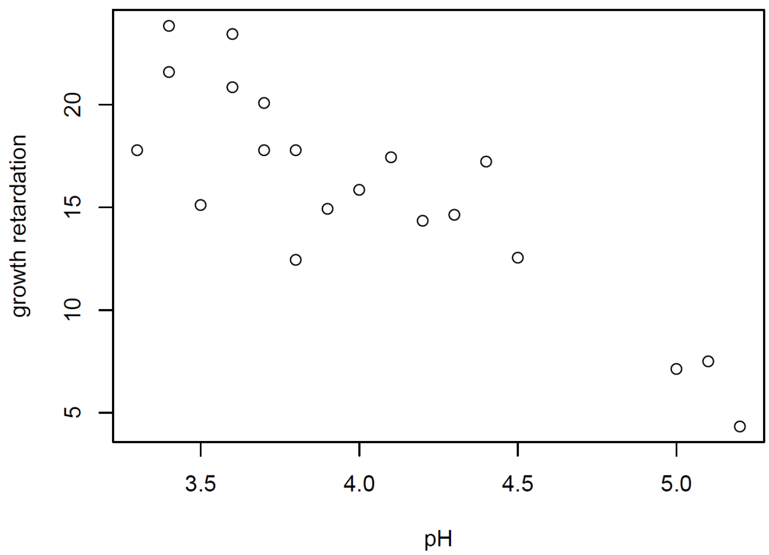
Note: *Throughout this exam, unless specifically stated otherwise, all notations/symbols and abbreviations used have their usual standard meanings, as used/defined in the lecture notes.*

Read the following information to answer **Questions 1–4**:

Forest scientists were interested in determining the impact of soil acidity on the retardation of tree growth. Data from $n = 20$ tree stands were collected. The measured variables were x = soil pH and y = index of growth retardation, and a simple linear regression model for y versus x was used. A scatterplot of the data and some summary statistics are given below.



$$\sum_{i=1}^{20} x_i = 80.5, \qquad \sum_{i=1}^{20} y_i = 316.84, \qquad \sum_{i=1}^{20}(x_i - \bar{x})(y_i - \bar{y}) = -49.022,$$

$$\sum_{i=1}^{20}(x_i - \bar{x})^2 = 6.2375, \qquad \sum_{i=1}^{20}(y_i - \bar{y})^2 = 518.6051, \qquad \sum_{i=1}^{20}(y_i - \hat{y}_i)^2 = 133.3295.$$

Use the information above to answer the **next four questions (Questions 1–4).**

1. The least-squares estimates of the intercept and slope parameters, respectively, are:

(a) $\hat{\beta}_0 = -7.86$ and $\hat{\beta}_1 = 47.48$.

(b) $\hat{\beta}_0 = 47.48$ and $\hat{\beta}_1 = -7.86$.

(c) $\hat{\beta}_0 = 15.84$ and $\hat{\beta}_1 = -7.86$.

(d) $\hat{\beta}_0 = -15.79$ and $\hat{\beta}_1 = 7.86$.

(e) $\hat{\beta}_0 = 213.16$ and $\hat{\beta}_1 = -49.022$.

Using the formulae from pg. 14 of Chapter 1A notes, the slope and intercept estimates are:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{20}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{20}(x_i-\bar{x})^2} = \frac{-49.022}{6.2375} = -7.86,\\ \hat{\beta}_0 &= \bar{y}-\hat{\beta}_1\bar{x} = \frac{316.84}{20}-(-7.86)\frac{80.5}{20} = 47.48.\end{aligned}$$

2. An estimate of σ, the error standard deviation, is given by:

(a) $518.6051/19$.

(b) $133.3295/18$.

(c) $\sqrt{133.3295/18}$.

(d) $\sqrt{518.6051/19}$.

(e) Cannot be determined from the information given.

Using the formula on pg. 18 of Chapter 1A notes, the required estimate of σ is:

$$\hat{\sigma} = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{133.3295}{18}}.$$

3. A 99% confidence interval for β_1 is given by:

(a) $\hat{\beta}_1 \pm 2.879(2.72)$.

(b) $\hat{\beta}_1 \pm 2.552(2.72)$.

(c) $\hat{\beta}_1 \pm 2.879(1.09)$.

(d) $\hat{\beta}_1 \pm 2.552(1.09)$.

(e) $\hat{\beta}_1 \pm 2.101(2.72)$.

The formula of $\hat{\sigma}$ is on pg. 18 of Chapter 1A notes. Using the formula from pg. 31 of Chapter 1A notes, the required 99% CI for β_1 is:

$$\begin{aligned}\hat{\beta}_1 \pm t_{n-2;\alpha/2}\hat{\sigma}_{\hat{\beta}_1} &= \hat{\beta}_1 \pm t_{18;0.005}\frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^{20}(x_i-\bar{x})^2}} = \hat{\beta}_1 \pm 2.879\frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^{20}(x_i-\bar{x})^2}}\\ &= \hat{\beta}_1 \pm 2.879\sqrt{\frac{133.3295/18}{6.2375}} = \hat{\beta}_1 \pm 2.879(1.09), \quad \text{where}\end{aligned}$$

we used $n = 20$, $\alpha = 0.01$ in this case, and the critical value $t_{n-2;\alpha/2} = t_{18;0.005} = 2.879$ from the t-table.

4. The proportion of variance in growth retardation that is explained by soil acidity is:

(a) 0.257.

(b) 0.743.

(c) 0.012.

(d) 0.611.

(e) 518.6.

We know (from pg. 25, Chapter 1A notes) that this proportion is R^2, or SSR/SST. We are given here that $SSE = \sum_{i=1}^{20}(y_i - \hat{y}_i)^2 = 133.3295$ and $SST = \sum_{i=1}^{20}(y_i - \bar{y})^2 = 518.6051$. So

$$R^2 = \frac{SSR}{SST} = \frac{(SST - SSE)}{SST} = \frac{(518.6051 - 133.3295)}{518.6051} = 0.743.$$

---

5. Having fitted a simple linear regression model to a set of n observations of (x_i, y_i) pairs, a good way to check if there is evidence of a nonlinear relationship between y and x is to:

(a) Determine the sum of squared of the residuals.

(b) Examine a plot of the residuals versus x_i.

(c) Examine a histogram of the residuals.

(d) Do both (b) and (c).

(e) Call Dr. Chakrabortty late at night for his opinion.

The correct answer is (b), as discussed on pg. 19-22 and pg. 46-48 of the notes. By definition, non-linearity is a type of relationship between y and x. Therefore, a histogram does not help identify non-linearity, since it discards information about how the residuals are related to x.

6. The lifetime Y (in thousands of holes drilled) for drill bits is related to the speed (x) of the drill through a simple linear regression model as follows:

$$Y = 6.0 - 0.017x + \epsilon \quad \text{with} \;\; 60 \leq x \leq 100, \quad \text{where}$$

ϵ has a Normal distribution with mean 0 and variance $\sigma^2 = 0.40$. Which of the following is the *highest* drill speed x such that at least 50% of drill bits used at that speed have lifetimes greater than 4.5? (**Hint:** for Normal distributions, the 50th percentile also equals the mean.)

(a) 60.

(b) 75.9.

(c) 88.2.

(d) 112.5.

(e) 136.1.

Using the linear model, for a given speed x (with $60 \leq x \leq 100$), Y is Normally distributed, specifically $Y \sim N(6-0.017x, 0.4)$ and we are being asked to find the *highest* speed x at which $P(Y > 4.5 \mid x) \geq 0.5$. So we need to find the highest x such that $P(6-0.017x+\epsilon > 4.5) \geq 0.5$, i.e. the highest speed x for which $P(\epsilon > 4.5 - 6 + 0.017x) \geq 0.5$, where $\epsilon \sim N(0, 0.4)$. Since the 50th percentile of any Normal distribution also equals the mean, and in this case ϵ has mean 0, this means $4.5 - 6 + 0.017x \leq 0$, and so $0.017x \leq 1.5$, i.e. $x \leq 1.5/0.017 = 88.2$. So, the required *highest* speed x for which the above condition is satisfied is $x = 88.2$.

Intuitively, since the lifetime Y at any speed x is Normally distributed, 50% of drill bits have lifetimes greater than the mean lifetime at that speed. So, in effect, we are being asked what is the speed x at which $6.0-0.017x = 4.5$. This is 88.2. At any speed higher than 88.2, the mean lifetime is lower than 4.5, and hence less than 50% of drill bits will have lifetimes greater than 4.5 at that speed. (Draw a picture! I drew such pictures several times in class.)

7. In the setting of Question **6**, what proportion of drill bits used at a speed $x = 75$ have lifetimes larger than 4?

(a) 0.0351.

(b) 0.1251.

(c) 0.5000.

(d) 0.8749.

(e) 0.9649.

Using the linear model, $Y \sim N(6.0 - 0.017x, 0.4)$ for any given speed x, and we are being asked to find $P(Y > 4)$ *given* $x = 75$. This probability is given by:

$$\begin{aligned}
P(Y > 4 \mid x = 75) &= P\left(Z > \frac{4 - 4.725}{\sqrt{0.40}}\right) = P(Z > -1.15), \quad \text{where } Z \sim N(0,1), \\
&= P(Z < 1.15) \qquad \text{(by symmetry of the } N(0,1) \text{ distribution)}, \\
&= 0.8749 \qquad \text{(using the Normal table)}.
\end{aligned}$$

8. Which of the following statements about simple linear regression is **not** correct?

(a) A 90% prediction interval for the response Y at a given $x = x_0$ is wider than a 90% confidence interval for the mean of Y at the same $x = x_0$.

(b) The coefficient of determination, R^2, always lies between 0 and 1.

(c) A 95% confidence interval for the mean of Y at a given value $x = x_0$ far from $\bar{x}$ will be narrower than a 95% confidence interval for the mean of Y at $x = \bar{x}$.

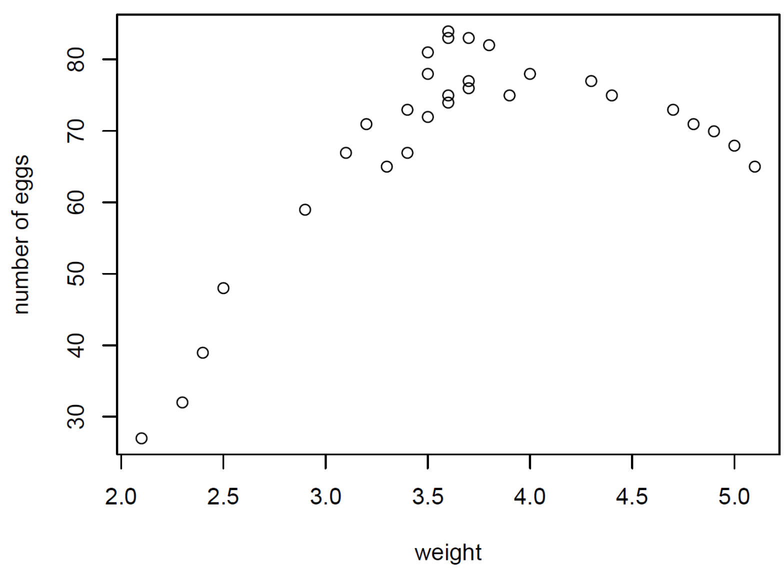
(d) The average increase in Y for a 1-unit increase in x is β_1 (i.e. the slope parameter).

(e) The parameter β_0 equals the y–intercept of the true regression line.

All options above are correct *except* for (c). The width of a 95% confidence interval for the mean of the response Y given x, in fact, gets wider the further the value of x is from $\bar{x}$. This follows directly from the relevant formulae given in pg. 34-37 of Chapter 1A notes. On the other hand, the fact that each of the other options are correct follows from the relevant portions of Chapter 1A notes and/or discussions in class.

Read the following information to answer **Questions 9–12**:

In a study of the reproductive success of grasshoppers, an entomologist collected a sample of $n = 30$ female grasshoppers. He recorded $y =$ the number of mature eggs produced and $x =$ the body weight (in grams) for each grasshopper. A scatterplot of the data is given below.



Polynomial models of degree 1 to 5 were fitted to these data. Information obtained from `R` is given below. Use this information to answer the **next four questions (Questions 9–12)**.

Polynomial degree	R^2	BIC	Estimated model
1	0.367	240.9	$27.80 + 11.24x$
2	0.942	172.5	$-155.71 + 116.01x - 14.30x^2$
3	0.942	175.8	$-171.48 + 130.32x - 18.43x^2 + 0.38x^3$
4	0.945	177.7	$112.68 - 210.76x + 130.48x^2 - 27.74x^3 + 1.94x^4$
5	0.945	181.1	$77.18 - 156.83x + 98.47x^2 - 18.46x^3 + 0.63x^4 + 0.073x^5$

$$\text{SST} = 6066.17, \quad \bar{y} = 68.83, \quad \bar{x} = 3.65.$$

9. Based on the information given, the best choice for the polynomial degree is:

(a) Degree 1, because it has the highest value of BIC.

(b) Degree 2, because it has the lowest value of BIC.

(c) Degree 3, because of Occam's razor.

(d) Either degree 4 or degree 5, because they have the largest R^2 values.

(e) Cannot be determined from the information given.

Besides the fact that the second degree model has the smallest BIC value, its R^2 is much larger than that of the first degree model and just marginally less than that of the 5th degree model. Therefore, both BIC and Occam's razor suggest using the degree 2 model.

10. Using the third degree model, an estimate of the *average* number of mature eggs produced by female grasshoppers weighing 3 grams is:

(a) 61.52.

(b) 62.76.

(c) 60.15.

(d) 63.87.

(e) 68.83.

Using the formula on pg. 70 of Chapter 1B notes, the required estimate of the mean response at $x = 3$ is simply the least squares estimate of the cubic model evaluated at $x = 3$, i.e.

$$-171.48 + 130.32(3) - 18.43(3^2) + 0.38(3^3) \;=\; 63.87.$$

11. The entomologist wants to predict the number of mature eggs that will be produced by a particular grasshopper (affectionately known as Jumpy) that weighs $x_0 = 4.5$ grams. Let

$\hat{\mu}(x)$ denote the kth degree polynomial that was finally chosen by the entomologist (according to some criteria of his liking). The following information was determined using R:

$$\hat{\mu}(4.5) = 76.78, \quad \hat{\sigma} = 3.605, \quad \text{estimated standard error of } \hat{\mu}(4.5)\text{: } SE_{\hat{\mu}(4.5)} = 0.988.$$

Then, we can be 95% sure that the number of eggs Jumpy produces will lie in the interval:

(a) $76.78 \pm t_{29-k;0.025}(0.988)$.

(b) $76.78 \pm t_{29-k;0.025}\sqrt{(3.605)^2 + (0.988)^2}$.

(c) $76.78 \pm 1.96(3.605)$.

(d) 76.78 ± 3.605.

(e) $3.605 \pm 1.96(76.78)$.

We're being asked for a *prediction* interval, since we're guessing the number of eggs produced by the particular grasshopper Jumpy. The formulae on pg. 71 of the Chapter 1B notes (and similar formulae on pg. 35-37 of Chapter 1A notes) imply that the prediction interval is:

$$76.78 \pm t_{29-k;0.025}\sqrt{\hat{\sigma}^2 + (SE_{\hat{\mu}(4.5)})^2}, \ \text{ i.e. } \ 76.78 \pm t_{29-k;0.025}\sqrt{(3.605)^2 + (0.988)^2},$$

where $SE_{\hat{\mu}(4.5)}$ is the estimated standard error of $\hat{\mu}(4.5)$. Hence, the correct answer is (b).

12. The estimate of the error variance, σ^2, based on the fourth degree model is given by:

(a) 0.945.

(b) 6066.17(0.945)/25.

(c) 6066.17/30.

(d) 6066.17(1 − 0.945)/25.

(e) Cannot be calculated from the information given.

Since we are given the R^2 value for each model, and also SST, we can find the error variance estimate $\hat{\sigma}^2$ for any such model. In particular, for the fourth degree model (i.e. $k = 4$), it is:

$$\hat{\sigma}^2 = \frac{SSE}{n-k-1} = \frac{SST - SSR}{30-4-1} = \frac{SST - SST(R^2)}{25} = \frac{6066.17 - 6066.17(0.945)}{25}.$$

---

13. A researcher has data $(x_1, y_1), \ldots, (x_n, y_n)$ and wants to fit the following linear model:

$$y_i = \beta_0 + \beta_1(1/x_i) + \epsilon_i, \quad i = 1, \ldots, n, \quad \text{where} \quad x_i \neq 0 \quad \text{and}$$

$\epsilon_i \sim N(0, \sigma^2)$ and $\epsilon_1, \ldots, \epsilon_n$ are independent. (You saw this model for the 'wind speed' data in class and in the homeworks.) Define $\bar{y} = \sum_{i=1}^n y_i/n$, $\bar{x} = \sum_{i=1}^n x_i/n$ and $\bar{x}_{inv} = \sum_{i=1}^n (1/x_i)/n$. Then, the correct expression for the least-squares estimate $\hat{\beta}_1$ of β_1 in this model is given by:

(a) $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}$.

(b) $\hat{\beta}_1 = \frac{\sum_{i=1}^n \{(1/x_i) - (1/\bar{x})\}(y_i - \bar{y})}{\sum_{i=1}^n \{(1/x_i) - (1/\bar{x})\}^2}$.

(c) $\hat{\beta}_1 = \frac{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv}\}(y_i - \bar{y})}{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv}\}^2}$.

(d) $\hat{\beta}_1 = \frac{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv}\}\{(1/y_i) - (1/\bar{y})\}}{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv}\}^2}$.

(e) $\hat{\beta}_1 = \frac{\sum_{i=1}^n \{(1/x_i) - (1/\bar{x})\}\{(1/y_i) - (1/\bar{y})\}}{\sum_{i=1}^n \{(1/x_i) - (1/\bar{x})\}^2}$.

You did **not** need to do any derivations whatsoever for answering this question, if you truly understood the calculations on pg. 12-14 of Chapter 1A notes (also in HW 1). Simply pretend that you have a simple linear regression problem of Y versus a 'transformed' covariate u where $u = 1/x$, the inverse of the 'original' covariate. Then, letting $\bar{u} = \sum_{i=1}^n u_i/n$, we note that $\bar{u} = \sum_{i=1}^n (1/x_i)/n = \bar{x}_{inv}$ (by notation). Using the formula of the least squares estimates from pg. 14 of Chapter 1A notes and adapting it accordingly to the situation here (involving the 'transformed' covariate u instead of the 'original' covariate x), we then have:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (u_i - \bar{u})(y_i - \bar{y})}{\sum_{i=1}^n (u_i - \bar{u})^2} = \frac{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv}\}(y_i - \bar{y})}{\sum_{i=1}^n \{(1/x_i) - \bar{x}_{inv})^2}.$$

14. In polynomial regression, the least squares estimates of the regression coefficients are the solution of a system of equations, the so-called 'normal equations'. These equations:

(a) Are non-linear and have an explicit solution.

(b) Are linear and do not have an explicit solution.

(c) Are non-linear and do not have an explicit solution.

(d) Are linear and have an explicit solution.

(e) Have befuddled budding statisticians for a good 150 years.

See pg. 51-52 of the Chapter 1B notes. We discussed these in detail multiple times in class.

15. Suppose that a response variable $Y > 0$ is such that $\log Y = \alpha + \beta x + \varepsilon$, where ε is a random variable, and α, β, x are constants (i.e. non-random), and the logarithm is taken to the base e. Then, the correct expressions for $E(Y)$ and $\mathrm{Var}(Y)$ are:

(a) $E(Y) = \alpha + \beta x + E(\varepsilon)$ and $\mathrm{Var}(Y) = \mathrm{Var}(\varepsilon)$.

(b) $E(Y) = e^{\alpha} e^{\beta x} e^{E(\varepsilon)}$ and $\mathrm{Var}(Y) = e^{\mathrm{Var}(\varepsilon)}$.

(c) $E(Y) = e^{\alpha} e^{\beta x} E(e^{\varepsilon})$ and $\mathrm{Var}(Y) = e^{2\alpha} e^{2\beta x} \mathrm{Var}(e^{\varepsilon})$.

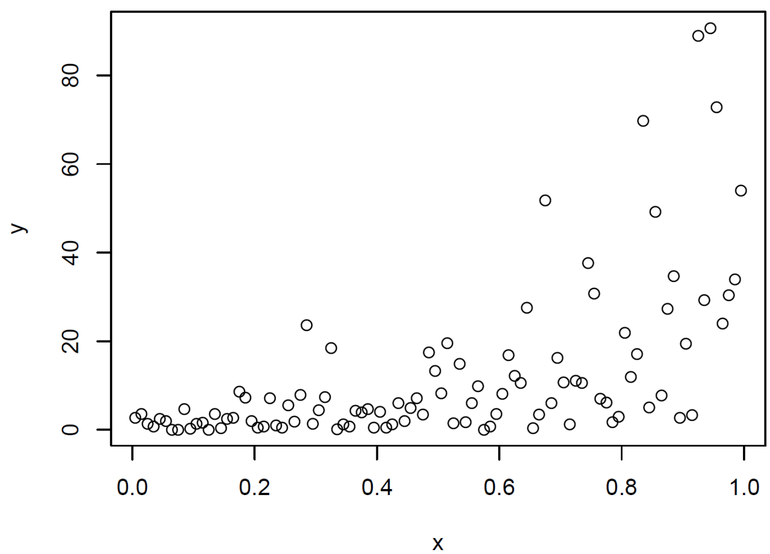
(d) $E(Y) = e^{\alpha}e^{\beta x}e^{E(\varepsilon)}$ and $\text{Var}(Y) = e^{2\alpha}e^{2\beta x}e^{\text{Var}(\varepsilon)}$.

(e) $E(Y) = e^{\alpha} + e^{\beta x} + E(e^{\varepsilon})$ and $\text{Var}(Y) = \text{Var}(e^{\varepsilon})$.

Since $\log Y = \alpha + \beta x + \varepsilon$, we have $Y = e^{\alpha+\beta x+\varepsilon} = e^{\alpha}e^{\beta x}e^{\varepsilon}$. Since α, β and x are constants (i.e. fixed and non-random), so are e^{α} and $e^{\beta x}$. Hence, using the basic properties: $E(cZ) = cE(Z)$ and $\text{Var}(cZ) = c^2\text{Var}(Z)$ for any random variable Z and any non-random constant c, we get:

$$\begin{aligned} E(Y) &= E(e^{\alpha}e^{\beta x}e^{\varepsilon}) = e^{\alpha}e^{\beta x}E(e^{\varepsilon}), \quad \text{and} \\ \text{Var}(Y) &= \text{Var}(e^{\alpha}e^{\beta x}e^{\varepsilon}) = (e^{\alpha}e^{\beta x})^2\text{Var}(e^{\varepsilon}) = e^{2\alpha}e^{2\beta x}\text{Var}(e^{\varepsilon}). \end{aligned}$$

16. Consider the following scatterplot of a regression data $(x_1, y_1), \ldots, (x_{100}, y_{100})$.



Which of the following is the *best* answer? (Remember: there is **no** partial credit!)

(a) It appears that the variance of the response variable Y is increasing with x.

(b) It appears that the variance of the response variable Y is decreasing with x.

(c) The variance of the response variable Y is approximately constant over all x.

(d) Using a log transformation of the response Y is a good idea in this case.

(e) Both (a) and (d) are correct.

See pg. 74-79 of Chapter 1B notes. We discussed these many times in class. Since this was straight out of the notes, you will **not** get any partial credit for picking out only one of these answers, i.e. options (a) and (d). You get credit only if you mark the *most* correct and inclusive option which is (e), as hinted at (along with a fair warning) in the question itself.

Table A.3 The Cumulative Distribution Function for the Standard Normal Distribution: Values of $\Phi(z)$ for Nonnegative z

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

Table A.4 Percentiles of the T Distribution

df	90%	95%	97.5%	99%	99.5%	99.9%
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.183	4.541	5.841	10.215
4	1.533	2.132	2.777	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.708	5.208
7	1.415	1.895	2.365	2.998	3.500	4.785
8	1.397	1.860	2.306	2.897	3.355	4.501
9	1.383	1.833	2.262	2.822	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.625	2.977	3.787
15	1.341	1.753	2.132	2.603	2.947	3.733
16	1.337	1.746	2.120	2.584	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.879	3.611
19	1.328	1.729	2.093	2.540	2.861	3.580
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.788	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.705	3.307
80	1.292	1.664	1.990	2.374	2.639	3.195
∞	1.282	1.645	1.960	2.326	2.576	3.090