8 Selected Answers

1.1
a) \( \text{diag}([3, 5], 2) \)

1.3
a) \( A \) must be square.
b) It’s unlikely any of the 20 pairs of matrices will be equal.

1.5 Almost any pair of square matrices will do the trick.

1.7
a) If \( A \) is \( m \)-by-\( n \) with \( m \neq n \), then the dimensions of \( A^T \) do not allow equality of \( A \) and \( A^T \).

1.8
a) In the long term, there are 4000 subscribing households and 6000 non-subscribers.

1.9 It is quite possible that none of your randomly generated matrices will be singular.

1.10
a) 6
b) 2-by-2, 2-by-2, 4-by-3, and 4-by-2 respectively

1.11
a) \( P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \)
8 Selected Answers

1.14 The matrix in letter (a) is in echelon form. Its 2\textsuperscript{nd}, 4\textsuperscript{th}, 5\textsuperscript{th} and 6\textsuperscript{th} columns are pivot columns with pivots 2, 3, 1 and 2 respectively.

The matrix in letter (c) is in echelon form. Its 1\textsuperscript{st} and 3\textsuperscript{rd} columns are pivot columns with pivots 1 and 3 respectively.

The matrices in letters (d) and (e) are in echelon form, both having pivot 1 in the only pivot column, column 1.

The matrices in letters (b) and (f) are not in echelon form.

1.15

a) (11, 3)
b) (4, 1, 3)

1.16

a) \[
\begin{align*}
3x_1 + 2x_2 &= 8 \\
x_1 + 5x_2 &= 7 \\
2x_1 + x_2 + 4x_3 &= -1
\end{align*}
\]
c) \[
\begin{align*}
4x_1 - 2x_2 + 3x_3 &= 4 \\
5x_1 + 2x_2 + 6x_3 &= -1
\end{align*}
\]

1.18

a) There is just one solution: \( x = (-1, -5/2, 2) \).
b) Solutions take the form \((3/2, 2, 0, -5/2) + t(-2, 0, 1, 0), t \in \mathbb{R} \).
c) There is no solution.

1.19 There is only one solution, namely \((7.5, -2.5, 2)\).

1.20

a) There is only one choice: \( c_1 = 4, c_2 = 3 \) and \( c_3 = 4 \).

C) There are infinitely many choices of the constants. They all take the form

\[
(c_1, c_2, c_3) = t(-3, 4, 1) + (1, -5, 0), \quad t \in \mathbb{R}.
\]
1.22 \text{null}(A) = \text{span}((-2, 1, 0, 0, 0), (-3, 0, 0, 0, 1))

1.23

a) Hint: Study the algorithm in simpleGE.

b) \((-10/3, 5/3, 2, 5/3)\)

1.24 Both systems have unique solutions: \(x_1 = (-1, 2, 1)\) and \(x_2 = (3, 1, -2)\).

1.26 Hint: If \(E_{ij}\) adds a multiple of row \(j\) to row \(i\), then the way to undo this is to subtract that same multiple.

1.27

a) (ii) \((A_1A_2)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -a_{21} & 1 & 0 \\ a_{32}a_{21} - a_{31} & -a_{32} & 1 \end{bmatrix}\) (iii) \(A_2A_1 = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} + a_{32}a_{21} & a_{32} & 1 \end{bmatrix}\)

(iv) \((A_2A_1)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -a_{21} & 1 & 0 \\ -a_{31} & -a_{32} & 1 \end{bmatrix}\)

b) (ii) \((A_1A_2A_3)^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{21} & 1 & 0 & 0 \\ a_{32}a_{21} - a_{31} & -a_{32} & 1 & 0 \\ a_{21}a_{42} + a_{31}a_{43} - a_{41} - a_{21}a_{32}a_{43} & a_{32}a_{43} - a_{42} & -a_{43} & 1 \end{bmatrix}\)

(iii) \(A_3A_2A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ a_{31} + a_{32}a_{21} & a_{32} & 1 & 0 \\ a_{41} + a_{31}a_{43} + a_{21}a_{42} + a_{21}a_{32}a_{43} & a_{42} + a_{32}a_{43} & a_{43} & 1 \end{bmatrix}\)

1.29

a) This interpolating linear polynomial would be of degree zero (i.e., be a constant function, one with zero slope) precisely when the two interpolation points have the same \(y\)-coordinate.

b) i. Given \(n\) points in the plane, no two of which share the same \(x\)-coordinate, there is a unique polynomial having degree at most \(n - 1\) that passes through these \(n\) points.
1.30 In this case our permutation matrix is just the identity. Thus, solving \( Ly = Pb \) is simply a matter of solving \( Ly = b \) using forward substitution. First, we solve

\[
Ly = Pb = \begin{bmatrix} -10 \\ -1 \\ -1 \end{bmatrix}
\]

via backward substitution to get \( y = (-10, -23/3, 10) \). Then we solve \( Ux = y \), which looks like

\[
\begin{align*}
6x_1 - 4x_2 + 5x_3 &= -10 \\
\frac{1}{3}x_2 + \frac{13}{3}x_3 &= \frac{-23}{3} \\
-5x_3 &= 10.
\end{align*}
\]

This leads to the solution \((2, 3, -2)\).

1.31 Solutions have the form \((-27, -7, 0) + z(-13, -5, 1), z \in \mathbb{R}\).

1.32

a) Here are sample results, taking \( \alpha = 0.75 \) (radians) as a “for instance”.

\[
\begin{align*}
octave-3.0.0:61> & \text{ alph } = .75; \\
octave-3.0.0:62> & \text{ A } = [\cos(\text{alph}) - \sin(\text{alph}); \sin(\text{alph}) \cos(\text{alph})]; \\
octave-3.0.0:63> & \text{ eig(A) ans } = \\
& 0.73169 + 0.68164i \\
& 0.73169 - 0.68164i
\end{align*}
\]

c) We have \( A = BC \) where

\[
B = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} \cos(0.4637) & -\sin(0.4637) \\ \sin(0.4637) & \cos(0.4637) \end{bmatrix}.
\]

1.33

a) For an \( n \)-by-\( n \) matrix, the eigenvalues are the roots of an \( n \)-th-degree polynomial. Such a polynomial has \( n \) roots, though some may repeated, which means it has at most \( n \) distinct roots. If \( A \) is 2-by-2, this means it can have at most one other root.
a) \( A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \), eigenpairs: \((-1, (1, 0)), (1, (0, 1))\)

b) \( A = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(\pi/4) & \sin(\pi/4) \\ -\sin(\pi/4) & \cos(\pi/4) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \)

(Note: Students can get this same matrix appealing to equation (1.6) with \( \alpha = 2\theta = \pi/2 \), or perhaps by appealing to part (b) of Exercise 1.13; these same observations are true of parts (c) and (d) below, with eigenpairs \((-1, (-1, 1)), (1, (1, 1))\)  (Any rescaling of the eigenvectors is permissible.)

1.35

a) \( A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \), eigenpairs: \((-1, i), (1, j), (1, k)\)

1.36

a) Three hours corresponds to a rotation through an angle of \( \pi/4 \). Thus,

\[
A = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) & 0 \\ \sin(\pi/4) & \cos(\pi/4) & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

1.37

a) octave-3.0.0:112> hPts = [0 0.5 0.5 4 4 0.5 0.5 5.5 5.5 0 0];
octave-3.0.0:113> hPts = [hPts; 0 0 4.5 4.5 5 5 5 7.5 7.5 8 8 0];
octave-3.0.0:114> hPts = [hPts; ones(1,11)];
octave-3.0.0:115> plot(hPts(1,:), hPts(2,:))
octave-3.0.0:116> axis("square")
octave-3.0.0:117> axis([-1 6 -1 9])

1.38

a) Hint: look at \( AS \) under the blocking

\[
AS = A \begin{bmatrix} S_1 & S_2 & \cdots & S_n \end{bmatrix},
\]
and SD under the blocking

\[ SD = \begin{bmatrix} S_1 & S_2 & \cdots & S_n \end{bmatrix} \]

\[
\begin{bmatrix}
  d_1 & 0 & 0 & \cdots & 0 \\
  0 & d_2 & 0 & \cdots & 0 \\
  0 & 0 & d_3 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \cdots & d_n
\end{bmatrix}
\]

b) Hint: If \( AS = SD \) and \( S \) is nonsingular, then \( A = SDS^{-1} \).

c) Hint: Recall that \( \det(AB) = \det(A) \det(B) \).

1.40 Any matrix with a row whose entries are all zero works.

1.41

a) \( \text{rank}(A) = 3, \text{nullity}(A) = 2 \)

b) A basis of \( \text{ran}(A) \) is \( B = \{(1, -1, 0, 1), (-2, 3, 1, 2), (2, -2, 4, 5)\} \)

c) One way: “Find a linear independent collection of vectors whose span is \( \text{ran}(A) \).”

1.42

a) The columns of \( A \) are linearly independent, since the lack of any nonzero element in \( \text{null}(A) \) indicates that the only linear combination of columns \( A_1, \ldots, A_n \) of \( A \) that yields the zero vector is

\((0)A_1 + (0)A_2 + \cdots + (0)A_n.\)

1.43 No.

1.44 Using a block interpretation of the product, we have

\[
uv = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} v = \begin{bmatrix} u_1 v \\ u_2 v \\ \vdots \\ u_m v \end{bmatrix}.
\]

This illustrates that each row of \( uv \) is a scalar multiple of \( v \). Thus, when we find the lead nonzero entry (the pivot) in row 1 of \( uv \) and then use \textbf{Elementary Operation} 3 to zero out
the entries in the column below this pivot, the result will be zero in every position \((i, j)\) of
the matrix with \(i > 1\).

1.46 \(\text{null}(A) = \{0\}\)

2.1 The additive identity (function) in \(\mathbb{F}^k(a, b)\) is the function that is constantly zero—
that is, \((x \mapsto 0)\). This function is differentiable (it is its own derivative) up to an order
you like (i.e., it has a \(k^{\text{th}}\) derivative for any choice of nonnegative integer \(k\)), and all of its
derivatives are continuous.

2.2

a) Hint: Explore what would happen if \([u, v]\) were linearly dependent.

b) \(3x - y + z = 0\)

c) Yes, this always happens. That is because, “containing the origin” means that \((0, 0, 0)\)
must satisfy the equation \(Ax + By + Cz = D\) or, equivalently, \(0 + 0 + 0 = D\).

2.3

a) Hint: Show that you can take two vectors in this plane and get a sum not in the
plane—i.e., that the plane is not closed under addition.

b) \(s(1, 0, 1) + t(0, 1, -1)\), \(s, \ t \in \mathbb{R}\).

c) Hint: Since \((1, -1, -1)\) is a normal vector to both planes, both should have equations
of the form \(x - y - z = D\).

2.4

c) Perhaps the easiest way to do this is in two steps. First, one shows that each \(u_j\) is
in \(\text{null}(H)\)—that is, show that \(Hu_j = 0\) for \(j = 1, \ldots, 4\). In doing so, we demonstrate
that the set \(S = \{u_1, u_2, u_3, u_4, w_1, w_2, w_3, w_4\}\) is a collection of vectors all of which are
from \(\text{null}(H)\). (Do you see why \(\text{span}(S) = \text{span}\{w_1, w_2, w_3, w_4\}\), and thereby that
\(\text{span}(S) = \text{null}(H)\)?) Next, we can form the matrix whose columns are the vectors
in \(S\)—that is, \(A = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 & w_1 & w_2 & w_3 & w_4 \end{bmatrix}\). Then by reducing \(A\) to
echelon form, we find that the first four columns are pivot columns while the last
four are free, thereby demonstrating that only the first four are needed as a basis for
\(\text{span}(S)\).

d) \((1, 0, 0, 1)\) (or 1001)
8 Selected Answers

2.5 \( \mathbb{R}^n \) is the set of all \( n \)-tuples whose elements are real numbers. That is,
\[
\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) | \text{each } x_i \in \mathbb{R}\}.
\]

2.6 Hint: Use the subspace test.

2.12

a) It is a basis.

c) It is not a basis.

2.14 \( \dim(\mathbb{Z}_2^n) = n \)

2.15 By supposition, \( \mathbf{u} \in \text{span}(S) \), so there exist real numbers \( a_1, \ldots, a_m \) for which
\[
\mathbf{u} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_m \mathbf{v}_m.
\]
We first show that \( \text{span}((\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u})) \) is a subset of \( \text{span}(S) \). Let \( \mathbf{w} \in \text{span}((\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u})) \). Then for some choice of real numbers \( b_1, \ldots, b_{m+1} \),
\[
\mathbf{w} = b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \cdots + b_m \mathbf{v}_m + b_{m+1} \mathbf{u}
\]
\[
= b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \cdots + b_m \mathbf{v}_m + b_{m+1}(a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_m \mathbf{v}_m)
\]
\[
= (b_1 + b_{m+1}a_1) \mathbf{v}_1 + (b_2 + b_{m+1}a_2) \mathbf{v}_2 + \cdots + (b_m + b_{m+1}a_m) \mathbf{v}_m,
\]
which shows \( \mathbf{w} \in \text{span}(S) \). Since \( \mathbf{w} \) is an arbitrary element of \( \text{span}((\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u})) \), we have that \( \text{span}((\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u})) \) is a subset of \( \text{span}(S) \).

Now we must show \( \text{span}(S) \) is a subset of \( \text{span}((\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u})) \). But this is much easier, because for any element \( \mathbf{w} \in \text{span}(S) \), there are real numbers \( b_1, \ldots, b_m \) for which
\[
\mathbf{w} = b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \cdots + b_m \mathbf{v}_m,
\]
which is already a linear combination of vectors in \( \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u}\} \).

2.19

a) In words, the left nullspace of \( \mathbf{A} \) is the set of vectors \( \mathbf{v} \in \mathbb{R}^m \) satisfying \( \mathbf{A}^T \mathbf{v} = \mathbf{0} \) or, equivalently, satisfying \( \mathbf{v}^T \mathbf{A} = 0^T \). Using set notation, this is
\[
\text{null}(\mathbf{A}^T) = \left\{ \mathbf{v} \in \mathbb{R}^m | \mathbf{A}^T \mathbf{v} = \mathbf{0} \right\}.
\]
2.20

a) \( \text{ran}(A) \) has basis \{ (2, 1, -3, 1), (3, 0, -5, 0) \}.
\( \text{ran}(A^T) \) has basis \{ (1, 0, 3, 1), (0, 1, -2, -1) \}.

For this particular matrix \( A \), the dimension of each subspace is 2.

3.2

a) As a start, we have
\[ \langle au + bv, w \rangle = \langle w, au + bv \rangle. \]

Now use linearity in the 2nd argument.

3.3 Suppose \( v \in \mathbb{R}^n \) satisfies \( \langle v, v \rangle = 0 \). Then
\[ 0 = \langle v, v \rangle = ||v||^2 = v_1^2 + v_2^2 + \cdots + v_n^2. \]
With the expression on the far right containing only perfect squares of real numbers, it follows that each \( v_j = 0 \). Since all of the components of \( v \) are zero, \( v \) is the zero vector.

3.4

a) Since \( 0 \) is orthogonal to every vector in \( \mathbb{R}^n \), the orthogonal complement of \( U \) contains every vector.

3.5

b) A basis for:
- \( \text{ran}(A^T) \): \{ (1, 0, -2), (0, 1, 1) \}
- \( \text{ran}(A) \): \{ (1, 2), (3, 4) \}
- (the standard basis is also appropriate here) \( \text{null}(A) \): \{ (2, -1, 1) \}
- \( \text{null}(A^T) = \{ 0 \} \): By convention a basis for the trivial subspace is the empty set \{ \}.

The vectors in \( \text{ran}(A) \) should be orthogonal to those in \( \text{null}(A^T) \); those in \( \text{ran}(A^T) \) should be orthogonal to those in \( \text{null}(A) \). Both of these relationships are realized.

3.6

a) The subspace \( U \) of \( \mathbb{R}^3 \) is precisely the row space of \( A = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \). We know \( U^\perp = \text{ran}(A^T)^\perp = \text{null}(A) \). Finding a basis for the latter, we get \{ (-1, 0, 1), (1, 1, 0) \}.
8 Selected Answers

3.7 Hint: \( \text{ran}(A^T)^\perp = \text{null}(A) \).

3.12

a) Hint: Any vector in \( \mathcal{U} \cap \mathcal{U}^\perp \) is going to be orthogonal to itself.

3.13 Suppose \( A \) is an \( m \)-by-\( n \) matrix. Theorem 3.1.6 says \( \text{null}(A) = \text{ran}(A^T)^\perp \). Taking the orthogonal complement of both sides, we get

\[
\text{null}(A)^\perp = (\text{ran}(A^T)^\perp)^\perp = \text{ran}(A^T),
\]

with this last equality holding because of Theorem 3.2.4 and the fact that \( \text{ran}(A^T) \) is a subspace of \( \mathbb{R}^n \).

3.16

b) Least-squares solutions form a line in the plane, and have the form \((-1/2, 0) + t(1/2, 1), \ t \in \mathbb{R} \).

d) The vector which is both a least-squares solution and lies in \( \text{ran}(A^T) \) is \((-2/5, 1/5)\).

e) \( p = (-1, -2, 1) \)

3.17

b) The least-squares solutions (of \( Ax = b \)) are

\[
(2.9 - t, 0.1, t) = t(-1, 0, 1) + (2.9, 0.1, 0), \quad t \in \mathbb{R}.
\]

3.18

a) We have \( A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \) and \( b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \).
d) We have
\[
\frac{\partial f}{\partial a_0} = \sum_{j=1}^{n} \frac{\partial}{\partial a_0} (y_j - a_0 - a_1 x_j)^2 = -2 \sum_{j=1}^{n} (y_j - a_0 - a_1 x_j)
\]
\[
= 2 \left( na_0 + a_1 \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} y_j \right), \quad \text{and}
\]
\[
\frac{\partial f}{\partial a_1} = -2 \sum_{j=1}^{n} x_j (y_j - a_0 - a_1 x_j)
\]
\[
= 2 \left( a_0 \sum_{j=1}^{n} x_j + a_1 \sum_{j=1}^{n} x_j^2 - \sum_{j=1}^{n} x_j y_j \right).
\]

Thus, the system of equations in \( a_0, a_1 \) which arises from setting these partial derivatives equal to zero is
\[
\begin{align*}
na_0 + a_1 (\sum_j x_j) - \sum_j y_j &= 0 \\
a_0 (\sum_j x_j) + a_1 (\sum_j x_j^2) - \sum_j x_j y_j &= 0
\end{align*}
\]
or
\[
\begin{align*}
na_0 + \alpha_x a_1 &= \alpha_y \\
\alpha_x a_0 + \beta_{xy} a_1 &= \beta_{xy},
\end{align*}
\]
where we have defined \( \alpha_x = \sum_j x_j, \alpha_y = \sum_j y_j, \beta_x = \sum_j x_j^2 \), and \( \beta_{xy} = \sum_j x_j y_j \).

Employing Cramer’s rule for \( a_1 \), we get
\[
a_1 = \frac{\begin{vmatrix} n & \alpha_y \\ \alpha_x & \beta_{xy} \end{vmatrix}}{\begin{vmatrix} n & \alpha_x \\ \alpha_x & \beta_x \end{vmatrix}} = \frac{n \beta_{xy} - \alpha_x \alpha_y}{n \beta_x - \alpha_x^2} = \frac{n \left( \sum_j x_j y_j \right) - \left( \sum_j y_j \left( \sum_j x_j \right) \right)}{n \left( \sum_j x_j^2 \right) - \left( \sum_j x_j \right)^2}.
\]

Substituting this answer into the top equation, we get
\[
a_0 = \frac{1}{n} (\alpha_y - a_1 \alpha_x) = \frac{1}{n} \left( \sum_j y_j - a_1 \sum_j x_j \right).
\]

4.2

a) Note that the distribution is bimodal.
8 Selected Answers

a) Hint: Treat separately the cases where the number of values is odd/even.

4.5
b) \( \bar{x}_3 = (2/3)\bar{x}_2 + (1/3)x_3 \)

4.8 Hint: What will be true of the relative sizes of these two quantities: the maximum value minus the third quartile, and the 1st quartile minus the minimum value?

4.11 Have all the values be the same. Then the variance is zero.

4.12
a) \( \hat{y} = \hat{x} + c \) (but be able to show it is true)

4.13
a) \( \hat{y} = c\hat{x} \)

4.14 To produce side-by-side boxplots, use R commands

```r
> data(barley)
> bwplot(yield ~ site, data = barley)
> bwplot(yield ~ year, data = barley)
> bwplot(yield ~ variety, data = barley)
```

There does not seem to be strong evidence here that \( yield \) was much affected by the \( variety \). Even though the “upper whisker” for 1931 extends higher than that for 1932, there is a great deal of overlap in the two distributions, casting much doubt about the effect \( year \) has on \( yield \). A similar observation may be made about the effect of \( site \) on \( yield \), though perhaps there is a significant difference between the \( yields \) in Grand Rapids vs. Waseca.

4.16 Hint: Drive down Surgeon B’s overall success rate by having her work on more “severe” cases than Surgeon A.

4.17 Concerning the mean, we may use the results of Exercises 4.12 and 4.13 (their (a) parts), to get

\[
\bar{z} = \frac{1}{s_x} - \frac{\bar{x}}{s_x} = 0 .
\]

4.18 You might try code like the following to see how small the sum of squares of relative errors is when using the one hundred meter time as a predictor of other times. There may be a different distance that does a better job of predicting all times.
4.19

b) approximately 823 runs

4.20

b) approximately 794 runs

4.21 \[ b = \frac{1}{||x||^2} \sum_{i=1}^{n} x_i y_i \]

4.22

a) (wins - losses) = 0.2142(runs - opponents’ runs)

4.24

a) We have
\[ y = \frac{b_0}{b_1 + x} \quad \Rightarrow \quad \frac{1}{y} = \frac{b_1 + x}{b_0} = \frac{b_1}{b_0} + \frac{1}{b_0} x. \]
This gives us \( g(y) = b'_0 + b'_1 h(x) \), with
\[ g(y) = \frac{1}{y}, \quad h(x) = x, \quad b'_0 = \frac{b_1}{b_0}, \quad \text{and} \quad b'_1 = \frac{1}{b_0}. \]

4.25 Linearization approach: \( b_0 = 195.802, b_1 = 0.0484; \) SSResid = 1920.747
Nonlinear LS-fit: \( b_0 = 212.684, b_1 = 0.06412; \) SSResid = 1195.449

4.26

e) There may be siblings included in the sample, exaggerating the representation of certain features not shared by the student body at large. Also, if there is a certain ethnic group (Dutch?) that typically has last names beginning with “A” more frequently than other ethnic groups at the college, then their representation in the sample may be out of proportion with their numbers at the college.
8 Selected Answers

4.28

b) We would expect 1 in 3 to be prime. In a sample of size 10, the resulting number 3.3 would not be a realizable by a single sample, so a single representative sample would likely have 3 (or possibly 4) prime numbers in it.

c) Answers will vary.

4.29

a) This one is appropriate.

4.30

a) There are 20.

4.31 One method which would qualify as a *simple* random sample is to assign from 1 to $N$ the verses in the Bible, then use a random number generator to choose 60 numbers at random from the list $1, \ldots, N$.

Of course, the problem asks for an alternative to this.

4.32 Hint: Use a weighted average.

4.36

a) The explanatory variable is level of pain medication before surgery, while the response variable is the level of pain after surgery. (The website lists some secondary response variables such as “time to first analgesic request”, “time to first oral intake”, “length of hospitalization”.)

4.37

a) The explanatory variable seems to be the categorical variable whose values indicate whether the plant was chilled overnight or not. The response variable is the carbon dioxide uptake rate.

4.39

b) One must be careful how one treats other human beings. At a minimum, one ought to know that a treatment is harmless before applying it to people. And, as long as smokers are self-selecting, there is no opportunity to randomize their placement into the experimental group.
5.1
a) HH, HT, TH, TT

5.2

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1</td>
<td>1,2</td>
<td>1,3</td>
<td>1,4</td>
<td>1,5</td>
<td>1,6</td>
</tr>
<tr>
<td>2,1</td>
<td>2,2</td>
<td>2,3</td>
<td>2,4</td>
<td>2,5</td>
<td>2,6</td>
</tr>
<tr>
<td>3,1</td>
<td>3,2</td>
<td>3,3</td>
<td>3,4</td>
<td>3,5</td>
<td>3,6</td>
</tr>
<tr>
<td>4,1</td>
<td>4,2</td>
<td>4,3</td>
<td>4,4</td>
<td>4,5</td>
<td>4,6</td>
</tr>
<tr>
<td>5,1</td>
<td>5,2</td>
<td>5,3</td>
<td>5,4</td>
<td>5,5</td>
<td>5,6</td>
</tr>
<tr>
<td>6,1</td>
<td>6,2</td>
<td>6,3</td>
<td>6,4</td>
<td>6,5</td>
<td>6,6</td>
</tr>
</tbody>
</table>

b) 1,6  2,5  3,4  4,3  5,2  6,1

5.3 Perhaps any of the numbers (given in inches) 48, 48.25, 48.5, 48.75, 49, \ldots, 81, 81.25, 81.5, 81.75, 82. There are perhaps occasional students outside this range, and in that instance it can be expanded. More significantly, though height is usually thought of as a continuous random variable, it is really not practical to measure it much more accurately than to the nearest quarter-inch.

5.5
b) The probability of 1 head is 1/32. The probability of 2 is 5/32.

5.6
b) The outcomes with exactly one head are easily listed, and there are 10 of them.

5.7
a) We would expect each face of the die to come up approximately 100 times.

b) Answers will vary.

5.8 Answers will vary.

5.9 Suggestion: One can list all of the possible ways in which 2 students can be selected from these 10. Using lower-case letters a–d for males and upper-case letters E–J for females, such a list would include outcomes like

ab, ac, ad, aE, aF, \ldots.
8 Selected Answers

5.10 Hint: One approach might be to simulate getting flushes in a particular suit. Perhaps create a vector of 1’s and 0’s via a command like

| > deck = c(replicate(13,c(1)), replicate(49,c(0)))

and then use commands like those found in Section A.6 of Appendix A. Another R command you may also find useful is prod, which works much like the sum command except that, instead of adding up all the components of a vector, it computes the product of those components.

5.11 a) Hint: Try something like

| > sum(sample(0:1, 1000, replace=T))

5.13 Hint: Start by adapting the hint from Exercise 5.10.

5.15

b) A reasonable estimate of this probability might be between 0.2 and 0.25.

5.17

a) 0.3484

5.18

a) 0.818

5.19 Without revealing the answers to parts (a) and (b), their difference is approximately 0.001254.

5.20

a) If $\pi$ is the probability of the outcome we are calling a “success”, then $(1 - \pi)$ is the probability of a “failure”. $x$ or fewer successes corresponds to $n - x$ or more failures. So, for $\pi = 0.6, 0.7, 0.8, 0.9$, then, $F(x; n, \pi) = 1 - F(n - x - 1; n, 1 - \pi)$ where, because the table would include entries for $1 - \pi = 0.1, \ldots, 0.5$, the value on the right-hand side is supplied.

5.21
a) Clearly it is the case that \( f_X(x) \geq 0 \) for all \( x \). Also, we have

\[
\int_{-\infty}^{\infty} f_X(x) \, dx = \int_{0}^{1} 2x \, dx = x^2 \bigg|_{0}^{1} = 1 .
\]

5.22
a) \( k = -3/32 \)
b) \( 1/2 \)

5.24 Let \( h = \alpha f + (1 - \alpha)g \). Then for each real \( x \),

\[
h(x) = \alpha f(x) + (1 - \alpha)g(x) \geq \alpha(0) + (1 - \alpha)(0) = 0 .
\]

Thus, the nonnegativity requirement for pdfs is satisfied by \( h \). The other requirement is satisfied as well, since

\[
\int_{-\infty}^{\infty} h(x) \, dx = \int_{-\infty}^{\infty} [\alpha f(x) + (1 - \alpha)g(x)] \, dx
\]

\[
= \alpha \int_{-\infty}^{\infty} f(x) \, dx + (1 - \alpha) \int_{-\infty}^{\infty} g(x) \, dx
\]

\[
= (\alpha)(1) + (1 - \alpha)(1) = 1 .
\]

5.25
b) \( 0.2 \)

5.26 \( \ln 2) / \lambda \)

5.27
a) \( 0.0198 \)
b) \( 0.9999922 \)

5.28
a) \( 0.1478562 \)
8 Selected Answers

b) 0.2369278

5.29 For \(X \sim \text{Unif}(a, b)\), we have

\[
E(X) = \int_{-\infty}^{\infty} x f_X(x) \, dx = \int_{a}^{b} \frac{x}{b-a} \, dx = \cdots = \text{the appropriate result.}
\]

5.30

b) 0.1096

5.31

a) With commands like

```r
> calKzoo = read.csv('http://www.calvin.edu/~stob/data/scores.csv')
> histogram(~time, data=calKzoo, breaks=10)
```

we get a histogram whose appearance is at least possibly suggestive of an exponential probability density function.

5.32 Hint: With so little guidance as to what kind of distribution \(X\) should have or what kind of transformation \(t(X)\) should be, one would have to think that, so long as you avoid transformations of the type discussed in Lemma 5.4.8, you will get the desired result.

5.33 Hint: Consider the random variable \(Y = X^2 - E(X)^2\).

5.36

a) approximately 2.5%

c) The top 25% include IQs exceeding 110.

5.37

a) The variance is 4/5.

5.38

a) \(\mu_X = 10, \text{Var}(X) = 100\)
b) \( \mu_X = \mu_X = 10, \text{Var}(X) = \text{Var}(X)/20 = 5 \)

6.1

a) 

b) We take \( \mu \) to be George’s average using his new ball. Then \( H_0: \mu = 211 \) and \( H_a: \mu \neq 211 \).

6.2

a) Note: The main issue is whether the alternative is one or two-sided.

b) It is about 0.043.

c) approximately 0.228

c) approximately 0.6

6.4

b) \( X \leq 34 \)

6.5

a) Note: The hypotheses should be statements the value of a parameter of the distribution.

b) If \( X \) is the wait time, then \( P(X = 2) = 0 \).

6.6

a) \((22.998, 23.050)\)

6.7

384

6.9

c) \((0.1518, 0.3210)\).
8 Selected Answers

6.10

a) For casein it is (282.64, 364.52).

c) Suggestion: Look at the graphical displays of the distributions broken down by feed.

6.12

b) The output speaks an “alternative hypothesis”, and the help for this command indicates one can include a switch indicating if a 1 or 2-sided alternative is desired. Evidently, the command t.test may be used for hypothesis testing as well as for constructing confidence intervals.

Many of us have been taught body temperature is 98.6°F. If this is so, then our sample has led to a 95% CI which did not capture the mean, which does happen on average about 1 in 20 times when a 95% CI is constructed. Another explanation might be that the 98.6 number is not accurate.

c) Hint: Address whether the population is normal (or, if not, whether it is still possible to consider X as approximately normal). You should also consider whether, in considering this sample as behaving “like a SRS”, they may be taken as instances of i.i.d. random variables.

6.13

a) We assume the speed of light is a fixed quantity c, and these measurements are likely distributed symmetrically around it. It is quite possible that there is a greater tendency to get readings close to c over readings far away, so it is possible that a normal distribution with mean c provides a good model. (836.72, 868.08).

6.15

b) A histogram of the data shows it is somewhat negatively skewed. With the size of the sample being n = 24, it is not skewed so badly as to make the resulting CI untrustworthy.

6.16

a) (−1.585, 6.494).

6.17
a) There appear to be differences in height between sopranos and altos and between sopranos and basses. This conclusion is based on 95% confidence intervals for the difference in means, noting that neither 95%-CI contains 0.

6.18
a) (RG) = 3.05 + 1.74 (HRG)

6.19
d) Hint: Look at a plot of residuals vs. age. Exercise 4.19 pointed to how one might obtain these residuals.

7.1
b) It is not smooth. (You still have to explain why.)

7.4
b) For the parametrization $r_1$, we have arc length

$$
\int_{-1}^{1} \sqrt{1 + 4t^2} \, dt = 2 \int_{0}^{1} \sqrt{1 + 4t^2} \, dt \approx 2.9579,
$$

this latter value computed using the applet.

7.6
a) $r(t) = (\cos t) \, i + (\sin t) \, j, \ t \in [0, 2\pi]$  
c) We might parametrize the lower side of the rectangle by

$$
r_1(t) = (t + 1) \, i + 3 \, j, \quad t \in [0, 2].
$$

Corresponding parametrizations for the other three sides (so that the entire rectangle is traversed counterclockwise) might be

$$
\begin{align*}
r_2(t) & = 3 \, i + (t + 1) \, j, \quad t \in [2, 7], \\
r_3(t) & = (10 - t) \, i + 8 \, j, \quad t \in [7, 9], \\
r_4(t) & = i + (17 - t) \, j, \quad t \in [9, 14].
\end{align*}
$$
8 Selected Answers

Then a piecewise smooth parametrization is given by

\[ r(t) = \begin{cases} 
  r_1(t), & t \in [0, 2], \\
  r_2(t), & t \in [2, 7], \\
  r_3(t), & t \in [7, 9], \\
  r_4(t), & t \in [9, 14]. 
\end{cases} \]

7.8

a) \( \mathbf{i} + \mathbf{j} + \mathbf{k} \)

b) Hint: In MATH 162 we learn that, given a normal vector \( a \mathbf{i} + b \mathbf{j} + c \mathbf{k} \) to a plane and a point \((x_0, y_0, z_0)\) on that plane, it has equation

\[ \langle (x - x_0) \mathbf{i} + (y - y_0) \mathbf{j} + (z - z_0) \mathbf{k}, a \mathbf{i} + b \mathbf{j} + c \mathbf{k} \rangle = 0, \]

or

\[ a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \]

7.9

a) The graph is a helical spiral. Its graph is depicted in Figure 11.4 of University Calculus, p. 664.

c) Hint: See the text immediately following Example 7.1.3.

7.10

c) Here the region lies between 0 and \( \pi \) along the \( s \)-axis, between 0 and \( 2\pi \) along the \( t \)-axis.