Solutions for Functions

Fn-1.1 (a) We know the domain and range of f. f is not an injection. Since no order is given for the domain, the attempt to specify f in one-line notation is meaningless (the ASCII order $+, <, >, ?$, is a possibility, but is unusual enough in this context that explicitly specifying it would be essential). If the attempt at specification makes any sense, it tells us that f is a surjection. We cannot give it in two-line form since we don't know the function.

(b) We know the domain and range of f and the domain has an implicit order. Thus the one-line notation specifies f. It is an injection but not a surjection. In two-line form it is $\begin{pmatrix} 1 & 2 & 3 \\ ? & < & + \end{pmatrix}$.

(c) This function is specified and is an injection. In one-line notation it would be (4,3,2), and, in two-line notation, $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \end{pmatrix}$.

Fn-1.2 (a) If f is an injection, then $|A| \le |B|$. **Solution**: Since f is an injection, every element of A maps to a different element of B. Thus B must have at least as many elements as A.

(b) If f is a surjection, then $|A| \ge |B|$. **Solution**: Since f is a surjection, every element of B is the image of at least one element of A. Thus A must have at least as many elements as B.

(c) If f is a bijection, then $|A| = |B|$. **Solution**: Combine the two previous results.

(d) If $|A| = |B|$, then f is an injection if and only if it is a surjection. **Solution**: Suppose that f is an injection and not a surjection. Then there is some $b \in B$ which is not the image of any element of A under f. Hence f is an injection from A to $B - \{b\}$. By (a), $|A| \le |B - \{b\}| < |B|$, contradicting $|A| = |B|$.
Now suppose that f is a surjection and not an injection. Then there are $a, a' \in A$ such that $f(a) = f(a')$. Consider the function f with domain restricted to $A - \{a'\}$. It is still a surjection to B and so by (b) $|B| \le |A - \{a'\}| < |A|$, contradicting $|A| = |B|$.

(e) If $|A| = |B|$, then f is a bijection if and only if it is an injection or it is a surjection. **Solution**: By the previous part, if f is either an injection or a surjection, then it is both, which is the definition of a bijection.

Fn-1.3 (a) Since ID numbers are unique and every student has one, this is a bijection.

(b) This is a function since each student is born exactly once. It is not an surjection since D includes dates that could not possibly be the birthday of any student; e.g., it includes yesterday's date. It is not an injection. Why? You may very well know of two people with the same birthday. If you don't, consider this. Most entering freshman are between 18 and 19 years of age. Consider the set F of those freshman and their possible birth dates. The maximum number of possible birth dates is $366 + 365$, which is smaller than the size of the set F. Thus, when we look a the function on F it is not injective.

(c) This is not a function. It is not defined for some dates because no student was born on that date. For example, D includes yesterday's date

(d) This is not a function because there are students whose GPAs are outside the range 2.0 to 3.5. (We cannot *prove* this without student record information, but we can be sure it is true.)

(e) We cannot *prove* that it is a function without gaining access to student records; however, we can be sure that it is a function since we can be sure that each of the 16 GPAs between 2.0 and 3.5 will have been obtained by many students. It is not a surjection since the codomain is larger than the domain. It is an injection since a student has only one GPA.

Fn-2.1 (a) For (1,5,7,8) (2,3) (4) (6): $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 3 & 2 & 4 & 7 & 6 & 8 & 1 \end{pmatrix}$ is the two-line form and (5,3,2,4,7,6,8,1) is the one-line form. (We'll omit the two-line form in the future since it is simply the one-line form with $1, 2, \ldots$ placed above it.) The inverse is (1,8,7,5) (2,3) (4) (6) in cycle form and (8,3,2,4,1,6,5,7) in one-line form.

(b) For $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 3 & 7 & 2 & 6 & 4 & 5 & 1 \end{pmatrix}$: The cycle form is (1,8) (2,3,7,5,6,4). Inverse: cycle form is (1,8) (2,4,6,5,7,3); one-line form is (8,4,2,6,7,5,3,1).

(c) For (5,4,3,2,1), which is in one-line form: The cycle form is (1,5) (2,4) (3). The permutation is its own inverse.

(d) (5,4,3,2,1), which is in cycle form: This is not the standard form for cycle form. Standard form is (1,5,4,3,2). The one-line form is (5,1,2,3,4). The inverse is (1,2,3,4,5) in cycle form and (2,3,4,5,1) in one-line form.

Fn-2.2 Write one entire set of interchanges as a permutation in cycle form. The interchanges can be written as (1,3), (1,4) and (2,3). Thus the entire set gives $1 \to 3 \to 2$, $2 \to 3$, $3 \to 1 \to 4$ and $4 \to 1$. In cycle form this is (1,2,3,4). Thus five applications takes 1 to 2.

Fn-2.3 (a) Imagine writing the permutation in cycle form. Look at the cycle containing 1, starting with 1. There are $n-1$ choices for the second element of the cycle AND then $n-2$ choices for the third element AND $\cdots$ AND $(n-k+1)$ choices for the kth element. Prove that the number of permutations in which the cycle generated by 1 has length n is $(n-1)!$: The answer is given by the Rule of Product and the above result with $k = n$.

(b) For how many permutations does the cycle generated by 1 have length k? We write the cycle containing 1 in cycle form as above AND then permute the remaining $n-k$ elements of $\underline{n}$ in any fashion. For the k long cycle containing 1, the above result gives $\frac{(n-1)!}{(n-k)!}$ choices. There are $(n-k)!$ permutations on a set of size $n-k$. Putting this all together using the Rule of Product, we get $(n-1)!$, a result which does not depend on k.

(c) Since 1 must belong to some cycle and the possible cycle lengths are $1, 2, \ldots, n$, summing the answer to (b) over $1 \le k \le n$ will count all permutations of $\underline{n}$ exactly once. In our case, the sum is $(n-1)! + \cdots + (n-1)! = n \times (n-1)! = n!$.

This problem has shown that if you pick a random element in a permutation of an n-set, then the length of the cycle it belongs to is equally likely to be any of the values from 1 to n.

Fn-3.1 (a) The domain and range of f are specified and f takes on exactly two distinct values. f is not an injection. Since we don't know the values f takes, f is not completely specified; however, it cannot be a surjection because it would have to take on all four values in its range.

(b) Since each block in the coimage has just one element, f is an injection. Since $|\text{Coimage(f)}| = 5 = |\text{range of } f|$, f is a surjection. Thus f is a bijection and, since the range and domain are the same, f is a permutation. In spite of all this, we don't know the function; for example, we don't know $f(1)$, but only that it differs from all other values of f.

(c) We know the domain and range of f. From $f^{-1}(2)$ and $f^{-1}(4)$, we can determine the values f takes on the union $f^{-1}(2) \cup f^{-1}(4) = \underline{5}$. Thus we know f completely. It is neither a surjection nor an injection.

(d) This function is a surjection, cannot be an injection and has no values specified.

(e) This specification is nonsense. Since the image is a subset of the range, it cannot have more than four elements.

(f) This specification is nonsense. The number of blocks in the coimage of f equals the number of elements in the image of f, which cannot exceed four.

Fn-3.2 (a) The coimage of a function is a partition of the domain with one block for each element of Image(f).

(b) You can argue this directly or apply the previous result. In the latter case, note that since Coimage(f) is a partition of A, $|\text{Coimage}(f)| = |A|$ if and only if each block of Coimage(f) contains just one element. On the other hand, f is an injection if and only if no two elements of A belong to the same block of Coimage(f).

(c) By the first part, this says that $|\text{Image}| = |B|$. Since Image(f) is a subset of B, it must equal B.

Fn-3.3 (a) The list is $321, 421, 431, 432, 521, 531, 532, 541, 542, 543$.

(b) The first number is $\binom{x_1-1}{3} + \binom{x_2-1}{2} + \binom{x_3-1}{1} + 1 = \binom{2}{3} + \binom{1}{2} + \binom{0}{1} + 1 = 1$. The last number is $\binom{4}{3} + \binom{3}{2} + \binom{2}{1} + 1 = 10$. The numbers $\binom{x_1-1}{3} + \binom{x_2-1}{2} + \binom{x_3-1}{1} + 1$ are, consecutively, $1, 2, \ldots 10$ and represent the positions of the corresponding strings $x_1x_2x_3$ in the list.

(c) The list is $123, 124, 125, 134, 135, 145, 234, 245, 345$.

(d) If, starting with the list of (c), you form the list $(6 - x_1)(6 - x_2)(6 - x_3)$, you get $543, 542, 541, 532, 531, 521, 432, 431, 421, 321$ which is the list of (a) in reverse order. Thus the formula of (b) gives the positions $\rho(x_x, x_2, x_3)$ in reverse order of the list (c). Subtract $11 - \rho(x_x, x_2, x_3)$ to get the position in forward order.

Fn-3.4 (a) The first distribution of balls to boxes corresponds to the strictly decreasing string 863. The next such string in lex order on all strictly decreasing strings of lengh 3 from $\underline{8}$ is 864. To get the corresponding distribution, place the three moveable box boundaries under positions 8, 6, and 4 and put balls under all other positions in $\underline{8}$. The predecessor to 863 is 862. The second distribution corresponds to 542. Its successor is 543, its predecessor is 541.

(b) The formula $p(x_1, x_2, x_3) = \binom{x_1-1}{3} + \binom{x_2-1}{2} + \binom{x_3-1}{1} + 1$ gives the position of the string $x_1x_2x_3$ in the list of decreasing strings of length three from $\underline{8}$. We solve the equation $p(x_1,x_2,x_3) = \binom{8}{3}/2 = 28$ for the variables x_1,x_2,x_3. Equivalently, find x_1,x_2,x_3 such that $\binom{x_1-1}{3}+\binom{x_2-1}{2}+\binom{x_3-1}{1} = 27$. First try to choose x_1-1 as large as possible so that $\binom{x_1-1}{3} \le 27$. A little checking gives $x_1-1=6$, with $\binom{x_1-1}{3} = \binom{6}{3} = 20$. Subtracting, $27-20=7$. Now choose x_2-1 as large as possible so that $\binom{x_1-1}{2} \le 7$. This gives $x_2-1=4$ with $\binom{x_2-1}{2} = \binom{4}{2} = 6$. Now subtract $7-6=1$ and choose $x_3-1=1$. Thus, $(x_1,x_2,x_3) = (7,5,2)$. The first element in the second half of the list is the next one in lex order after 752 which is 753. The corresponding distributions of ball into boxes can be obtained in the usual way.

Fn-3.5 (a) $2,2,3,3$ is not a restricted growth (RG) function because it doesn't start with 1.
$1,2,3,3,2,1$ is a restricted growth function. It starts with 1 and the first occurrence of each integer is exactly one greater than the maximum of all previous integers.
$1,1,1,3,3$ is not an RG function. The first occurrence of 3 is *two* greater than the max of all previous integers.
$1,2,3,1$ is an RG function.

(b) We list the blocks $f^{-1}(i)$ in order of i. Observe that all partitions of 4 occur exactly once as coimages of the RG functions.

$1111 \to \{1,2,3,4\}$	$1112 \to \{1,2,3\},\{4\}$	$1121 \to \{1,2,4\},\{3\}$
$1122 \to \{1,2\},\{3,4\}$	$1123 \to \{1,2\},\{3\},\{4\}$	$1211 \to \{1,3,4\},\{2\}$
$1212 \to \{1,3\},\{2,4\}$	$1213 \to \{1,3\},\{2\},\{4\}$	$1221 \to \{1,4\},\{2,3\}$
$1222 \to \{1\},\{2,3,4\}$	$1223 \to \{1\},\{2,3\},\{4\}$	$1231 \to \{1,4\},\{2\},\{3\}$
$1232 \to \{1\},\{2,4\},\{3\}$	$1233 \to \{1\},\{2\},\{3,4\}$	$1234 \to \{1\},\{2\},\{3\},\{4\}$

(c) 11111, 11112, 11121, 11122, 11123 $\to \{\{1,2,3\},\{4\},\{5\}\}$
11211, 11212, 11213, 11221, 11222 $\to \{\{1,2\},\{3,4,5\}\}$
11223, 11231, 11232, 11233, 11234 $\to \{\{1,2\},\{3\},\{4\},\{5\}\}$

Fn-4.1

$h_{X,Y}$	0	1	2	3	4	f_X
0	1/16	0	0	0	0	1/16
1	0	4/16	0	0	0	4/16
2	0	3/16	3/16	0	0	6/16
3	0	0	2/16	2/16	0	4/16
4	0	0	0	0	1/16	1/16
f_Y	1/16	7/16	5/16	2/16	1/16	

The row index is X and the column index is Y.

$E(X)=2$, $\mathrm{Var}(X)=\sigma_X=1$ $\quad E(Y)==1.69$, $\mathrm{Var}(Y)=0.96$, $\sigma_Y=0.98$

(c) $\mathrm{Cov}(X,Y)=0.87$

(d) $\rho(X,Y)=0.87/(1)(0.98)=+0.89$ Since the correlation is close to 1, X and Y move up and down together. In fact, you can see from the table for the joint distribution that X and Y are often equal.

Fn-4.2 (a) You should be able to supply reasons for each of the following steps

$$\begin{aligned}
\mathrm{Cov}(aX+bY,aX-bY) &= E[(aX+bY)(aX-bY)] - E[(aX+bY)]E[(aX-bY)] \\
&= E[a^2X^2-b^2Y^2] - [aE(X)-bE(Y)][aE(X)+bE(Y)] \\
&= E[a^2X^2-b^2Y^2] - [a^2E^2(X)-b^2E^2(Y)] \\
&= a^2[E(X^2)-E^2(X)] - b^2[E(Y^2)-E^2(Y)] \\
&= a^2\mathrm{Var}(X) - b^2\mathrm{Var}(Y)
\end{aligned}$$

Alternatively, using the bilinear and symmetric properties of Cov:

$$\begin{aligned}\mathrm{Cov}(aX+bY,aX-bY) &= a^2\mathrm{Cov}(X,X) - ab\mathrm{Cov}(X,Y) + ba\mathrm{Cov}(Y,X) + b^2\mathrm{Cov}(Y,Y)\\ &= a^2\mathrm{Var}(X) - b^2\mathrm{Var}(Y)\end{aligned}$$

(b) Here is the calculation:

$$\begin{aligned}\mathrm{Var}[(aX+bY)(aX-bY)] &= \mathrm{Var}[a^2X^2 - b^2Y^2)]\\ &= a^4\mathrm{Var}(X^2) - 2a^2b^2\mathrm{Cov}(X^2,Y^2) + b^4\mathrm{Var}(Y^2)\end{aligned}$$

Fn-4.3 We begin our calculations with no assumptions about the distribution for X. Expand the argument of the expectation and then use linearity of expectation to obtain.

$$E((aX+b)^2) = E(a^2X^2 + 2abX + b^2)) = a^2E(X^2) + 2abE(X) + b^2\,.$$

(The last term comes from the fact that $E(b^2) = b^2$ since b^2 is a constant.) By definition, $\mathrm{Var}(X) + (E(X))^2 = E(X^2)$. Thus

$$E((aX+b)^2) = a^2\Big(\mathrm{Var}(X) + (E(X))^2\Big) + 2abE(X) + b^2\,.$$

With a little algebra this becomes,

$$E((aX+b)^2) = a^2\mathrm{Var}(X) + (aE(X)+b)^2\,.$$

Specializing to the particular distributions for parts (a) and (b), we have the following.

(a) $E((aX+b)^2) = a^2np(1-p) + (anp+b)^2$.

(b) $E((aX+b)^2) = a^2\lambda + (a\lambda+b)^2$.

Fn-4.4 We make the dubious assumption that the misprints are independent of one another. (This would not be the case if the person preparing the book was more careless at some times than at others.)

Focus your attention on page 8. Go one by one through the misprints m_1, m_2, …, m_{200} asking the question, “Is misprint m_i on page 8?”

By the assumptions of the problem, the probability that the answer is “yes” for each m_i is $1/100$. Thus, we are dealing with the binomial distribution $b(k;200,1/100)$. The probability of there being less than four misprints on page 8 is

$$\sum_{k=0}^{3} b(k;200,1/100) = \sum_{k=0}^{3}\binom{200}{k}(1/100)^k(99/100)^{200-k}.$$

Using a calculator, we find the sum to be 0.858034.

Using the Poisson approximation, we set $\lambda = np = 2$ and compute the easier sum

$$e^{-2}2^0/0! + e^{-2}2^1/1! + e^{-2}2^2/2! + e^{-2}2^3/3!,$$

which is 0.857123 according to our calculator.

Solutions for Functions

Fn-4.5 From the definition of Z and the independence of X and Y, Tchebycheff's inequality states that

$$P(|Z - aE(X) - bE(y)| \geq \epsilon) \leq \frac{\mathrm{Var}(X) + \mathrm{Var}(Y)}{\epsilon^2}.$$

Applying this to the two parts (a) and (b), we get

(a) $P(|Z - a\gamma - b\delta| \geq \epsilon)\ leq \frac{\gamma + \delta}{\epsilon^2}$.

(b) $P(|Z - anr - bns \geq \epsilon) \leq \frac{nr(1-r) + ns(1-s)}{\epsilon^2}$.

Fn-4.6 We are dealing with $b(k; 1000, 1/10)$. The mean is $np = 100$ and the variance is $npq = 90$. The standard deviation is thus, 9.49. The exact solution is

$$\sum_{k=85}^{115} b(k; 1000, 1/10) = \sum_{k=85}^{115} \binom{1000}{k} (1/10)^k (9/10)^{1000-k}.$$

Using a computer with multi-precision arithmetic, the exact answer is 0.898. To apply the normal distribution, we would compute the probability of the event $[100, 115.5]$ using the normal distribution with mean 100 and standard deviation 9.49. In terms of the standard normal distribution, we compute the probability of the event $[0, (115.5 - 100)/9.49] = [0, 1.63]$. This probability is 0.4484. We double this to get the approximate answer: 0.897.

Fn-4.7 We have

$$\begin{aligned} E(X) &= E((1/n)(X_1 + \cdots + X_n)) = (1/n)E(X_1 + \cdots + X_n) \\ &= (1/n)(E(X_1) + \cdots + E(X_n)) = (1/n)(\mu + \cdots + \mu) = \mu \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Var}((1/n)(X + 1 + \cdots + X_n)) = (1/n)^2\, \mathrm{Var}(X + 1 + \cdots + X_n) \\ &= (1/n)^2(\mathrm{Var}(X+1) + \cdots + \mathrm{Var}(X_n)) = (1/n)^2(n\sigma^2) = \sigma^2/n. \end{aligned}$$

Since X has mean μ, it is a reasonable approximation to μ. Of course, it's important to know something about the accuracy.

(c) Since Var(X)$= \sigma^2/n$, we have $\sigma_X = \sigma/\sqrt{n}$. If we change from n to N, σ_X changes to $\sigma/\sqrt{N}$. Since we want to improve accuracy by a factor of 10, we want to have $\sigma/\sqrt{N} = (1/10)(\sigma/\sqrt{n})$. After some algebra, this gives us $N = 100n$. In other words we need to do 100 times as many measurements!