1. Let $B$ denote the P-value for $>;$ let $C$ denote the P-value for $<$; and let $D$ denote the P-value for $\neq$.

The first thing to note is that

$$x = \hat{p}_1 - \hat{p}_2 = 0.35 - 0.30 = 0.05,$$

is a positive number. This implies that

$$B \leq 0.5000 < C,$$

and that

$$D = 2B.$$

Also, recall that $B + C$ must be greater than 1. We now argue by contradiction.

- I start by determining which P-value is 0.8326. This cannot be $B$ because it exceeds 0.5000.
  If $D = 0.8326$, then
  $$B = D/2 = 0.4163.$$  
  By the process of elimination,
  $$C = 0.2454.$$  
  Thus, $B + C < 1$, which is impossible.  
  Thus, $C = 0.8326$.

- If $D = 0.2454$, then
  $$B = D/2 = 0.1227.$$  
  From above,
  $$C = 0.8326.$$  
  Thus, $B + C < 1$, which is impossible.  
  Thus, $B = 0.2454$ and
  $$D = 2B = 0.4908.$$

2. The first thing to note is that

$$x = \hat{p}_1 - \hat{p}_2 = 0.52 - 0.60 = -0.08,$$

is a negative number. Define the following numbers:

$$B = P(X \leq -0.08); C = P(X \geq -0.08);$$

and $D = P(X \geq +0.08)$.

(It might help to draw a picture on the number line.) Note that because we cannot assume symmetry, $B$ and $D$ need not be equal. We do know, however, that $B + C > 1$ and $C > D$. With this notation, the P-values are given in the table below.

<table>
<thead>
<tr>
<th>Alternative</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&gt;$</td>
<td>$C$</td>
</tr>
<tr>
<td>$&lt;$</td>
<td>$B$</td>
</tr>
<tr>
<td>$\neq$</td>
<td>$B + D$</td>
</tr>
</tbody>
</table>

Let’s consider the P-value 0.2931.

- Suppose that $C = 0.2931$. Then
  $$B + C > 1 \text{ implies } B > 0.7069$$
  and $B + D$ must equal 0.5242, which is impossible because it is smaller than $B$.

- Suppose that $B + D = 0.2931$. Then
  $$B \leq 0.2931$$
  making $C = 0.5242$

and $B + C < 1$.

Thus, $B = 0.2931$. $C$ cannot equal 0.5242 because this would make $B + C < 1$. Thus, $B + D = 0.5242$ and $C$ is unknown.
3. (a) The completed table is below, with the fractions not reduced.

<table>
<thead>
<tr>
<th>$X_1$</th>
<th>1</th>
<th>3</th>
<th>5.5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>3/30</td>
<td>2/30</td>
<td>5/30</td>
</tr>
<tr>
<td>3</td>
<td>3/30</td>
<td>6/30</td>
<td>6/30</td>
<td>15/30</td>
</tr>
<tr>
<td>5.5</td>
<td>2/30</td>
<td>6/30</td>
<td>2/30</td>
<td>10/30</td>
</tr>
<tr>
<td>Total</td>
<td>5/30</td>
<td>15/30</td>
<td>10/30</td>
<td>1</td>
</tr>
</tbody>
</table>

For example, the probability of the cell (3,3) is $3/6$ (the probability the first card will be a '3') multiplied by $2/5$ (the probability the second card will be a '3' after the first card has been a '3').

(b) First, we note that for the sum to equal 8.5, either cell (3, 5.5) or (5.5, 3) must occur. Reading from the table above, this probability is $6/30 + 6/30 = 12/30$.

(c) First, we note that for the cards to be equal, either cell (1, 1), (3, 3) or (5.5, 5.5) must occur. Reading from the table above, this probability is $0 + 6/30 + 2/30 = 8/30$.

4. (a) Both critical regions have the ≥ sign; thus, the alternative for both tests is >.

(b) This is simply

$$P(U \geq 13) = 3/36 + 5/36 = 8/36.$$  

(c) This is simply

$$P(R_1 \geq 13) = 4/36 + 5/36 = 9/36.$$  

(d) Either both tests reject or both tests fail to reject. Thus, the probability is $24/36 + 5/36 = 29/36$.

5. (a) The cell count described is ‘b;’ thus, the count is 100.

(b) The cell count described is ‘d;’ thus, the count is 700.

(c) This one is a bit tricky. Because I am interested in only one power, and not a comparison of powers, I use the nearly certain interval from Chapter 4. $r_R = 900/2500 = 0.360$ and the nearly certain interval is $0.360 \pm 3 \sqrt{0.36(0.64)/2500} = 0.360 \pm 0.029$.

(d) The nearly certain interval is $200 - 100/2500 \pm (3/2500)\sqrt{100 + 200} = 0.040 \pm 0.021$.

6. (a) You are given the sequence 111100. The other four sequences are:

111101, 011110, 101111 and 001111.

(b) There are five sequences that yield ($V = 4$). Thus, $P(V = 4)$ is the sum of the probabilities of these five sequences:

$$3p^4q^2 + 2p^5q.$$  

7. (a) By counting,

$$x = 27, r = 25, v = 5 \text{ and } w = 3.$$  

(b) First,

$$c = 2(27)(45 - 27) = 972.$$  

Thus,

$$\mu = 1 + (972/45) = 22.6 \text{ and }$$  

$$\sigma = \sqrt{972(972 - 45)/45(45)(44)} = 3.18.$$
For the runs test and the alternative $>$, the P-value is $P(R \geq r) = P(R \geq 25)$. We are going to use a Normal curve to approximate this probability. We use the Normal curve that matches the distribution of $R$ on mean and variance, both of which we found in part (b). Thus, we enter $A = 22.6$ and $B = 3.18$.

We want the area to the right, so we enter a number for C and, because we want the continuity correction, we replace 25 by 24.5.

Similar to our answer for (c), we enter the same values for A and B and enter 25.5 for D.

Because the mean of the Normal curve, 22.6, is smaller than both 24.5 and 25.5, it follows that $P_1 < 0.5000 < P_2$.

The Normal curve is symmetric; hence, $P_3 = 2P_1$.

Let $B$ denote the P-value for $>$; let $C$ denote the P-value for $<$; and let $D$ denote the P-value for $\neq$. We know that $B + C > 1$ and that, by symmetry, $D$ is twice the smaller of $B$ and $C$.

First,
\[ x = 0.30 - 0.40 = -0.10 \]

is a negative number. Thus,
\[ C < 0.5000 < B \text{ and } D = 2C. \]

Let’s start with 0.1055. This cannot be $B$ because it is smaller than 0.5000. If it were $D$, then $C = D/2 = 0.0528$ and $B + C$ would be smaller than 1 for either of the possible values for $B$. Thus, $C = 0.1055$ and $D = 2C = 0.2110$.

If $B = 0.8822$, then $B + C < 1$; thus, $B = 0.9411$.

(a) We need to rewrite the event $(X_1 = X_2)$ in terms of cells in the table. It is $(X_1 = X_2 = 2)$ or $(X_1 = X_2 = 3)$.

Thus, its probability is
\[ 0.06 + 0.16 = 0.22. \]

(b) We need to write the event $(X_1 \times X_2 > 5)$ in terms of cells in the table. It consists of the cells (2,3), (3,2), (3,3), (4,2) and (4,3). Thus, its probability is
\[ 0.11 + 0.090.16 + 0.15 + 0.23 = 0.74. \]

(a) This is a power analysis which implies that the correct decision is to reject the null. Thus, the cell of interest is $R_1$ rejects and $U$ fails to reject; its count is 50.

(b) The cell of interest is $R_1$ rejects and $U$ rejects; its count is 650.

This one is a bit tricky. Because I am interested in only one power, and not a comparison of powers, I use the nearly certain interval from Chapter 4. $r_U = 900/2000 = 0.450$ and the nearly certain interval is
\[ 0.450 \pm 3 \sqrt{0.450(0.55)}/2000 =
\]
\[ 0.450 \pm 0.033. \]

The nearly certain interval is
\[ (250 - 50)/2000 \pm 3/2000 \sqrt{50 + 250} =
\]
\[ 0.100 \pm 0.026. \]

I will refer to each interval by its lower bound, suppressing the decimal point for (my) ease; i.e., they are 519, 528, 536 and 544.

First, we calculate the four centers. We get 0.631 twice and two other numbers. The two approximate intervals must have the same center ($\bar{p}$); thus, 528 and 544 are the approximate intervals.
Interval 544 is narrower than interval 528; thus, the former is 90% and the latter is 95%.

Finally, 519 and 536 are the exact intervals. The former is wider; thus, it is 95% and the latter is 90%.

12. (a) Reading from the table, the CI is 
\[ [0.108, 0.566] \]

(b) The CI contains \( p = 0.600 \) if, and only if, \( x \) equals 4, 5, 6, 7 or 8.

(c) The lower bound of the CI exceeds \( p = 0.300 \) if, and only if, \( x \) equals 6, 7, 8, 9 or 10.

(d) The upper bound of the CI is smaller than \( p = 0.700 \) if, and only if, \( x \) equals 0, 1, 2, 3 or 4.

13. It is important to go back in time before data are collected to set up the problem. We plan to observe \( X \) which has a Poisson distribution with parameter \( \theta \). For later in the problem, we note that 
\[ \theta = 15\lambda, \]
where \( \lambda \) is the rate per hour.

The data yield \( X = 840 \), making the 95% CI for \( \theta \):
\[ 840 \pm 1.96\sqrt{840} = 840 \pm 56.81. \]

Next, we divide thru by 15 to convert to \( \lambda \), obtaining:
\[ 56.00 \pm 3.79. \]

14. First, we calculate \( r' = 12/15 = 0.8 \). The 90% PI is:
\[ 0.8(840) \pm 1.645\sqrt{0.8(840)(1.8)} = 672 \pm 57.21. \]

15. As with problem 3, it helps to go back in time to before the data are collected. Vince plans to observe \( X \) which has a Poisson distribution with parameter \( \theta \), where 
\[ \theta = 3\lambda, \]
where \( \lambda \) is the rate per minute.

Vince observes \( X = 4 \) and the website gives Vince an upper bound, call it \( b \), for \( \theta \):
\[ \theta \leq b \text{ or } \lambda \leq b/3. \]

We are given that 
\[ b/3 = 2.5574; \text{ thus, } b = 7.6722. \]

The inequality
\[ \theta \leq 7.6722 \text{ becomes } 700p \leq 7.6722, \]
which gives 
\[ p \leq 0.01096. \]

16. (a) In order to find the critical region, we assume that the null hypothesis is correct; hence, we will obtain our probabilities from the column for \( p = 0.25 \).

The alternative is \( > \); thus, the critical region will have the form \( (X \geq c) \). We must determine the value of \( c \).

Per the instructions in the problem, we need to find \( c \) so that 
\[ P(X \geq c) < 0.10, \]
but as close to 0.10 as possible. Using trial-and-error, I begin with \( c = 6 \). From the table,
\[ P(X \geq 6) = 0.0917 + 0.0393 + \ldots. \]

My guess can’t be correct because the sum of the first two terms exceeds 0.10. Next, I try \( c = 7 \) and get 
\[ P(X \geq 7) = 0.0393 + 0.0131 + 0.0034 + 0.0007 + 0.0001 = 0.0566. \]

Thus, the critical region is \( (X \geq 7) \).

(b) From (a), \( \alpha = 0.0566 \).

(c) The null hypothesis is that \( p = 0.20 \); thus, we use this column to obtain probabilities. The value of \( \alpha \) is 
\[ P(X \geq 6) = 0.0430 + 0.0138 + 0.0035 + 0.0007 + 0.0001 = 0.0611. \]
(d) We want the power for \( p = 0.30 \); thus, we use this column to obtain probabilities. The value of the power is

\[
P(X \geq 6) = 0.1472 + 0.0811 + 0.0348 + 0.0116 + 0.0030 + 0.0006 + 0.0001 = 0.2784.
\]

17. The approximate 95% confidence interval for \( \theta \) is

\[
x \pm 1.96 \sqrt{x} = 960 \pm 1.96 \sqrt{960} = 960 \pm 60.72.
\]

We can infer that

\[
\theta = 12 \lambda.
\]

Thus,

\[
960/12 = 80 \pm 5.06
\]

is the approximate 95% confidence interval for \( \lambda \).

18. First, \( r' = 9/12 = 0.75 \). The 90% prediction interval is

\[
r'x \pm 1.645 \sqrt{r'x(1 + r')} = 0.75(960) \pm 1.645 \sqrt{0.75(960)(1.75)} = 720 \pm 58.4.
\]

19. Let \( \lambda \) denote the rate per minute. The website gives Vince the upper bound for \( \theta = 10 \lambda \).

You are told that the upper bound for \( \lambda \) is 0.7906; thus, the upper bound for \( \theta \) is 10(0.7906) = 7.906. Written as an inequality, we have

\[
\theta \leq 7.906.
\]

We want the upper bound for \( p \). We obtain this by using the Poisson to approximate the binomial. We replace \( \theta \) by \( np \) and get the following approximate upper bound

\[
np \leq 7.906.
\]

Noting that \( n = 800 \), we solve for \( p \) and get

\[
p \leq 0.009882.
\]

Note this approximation works only because Sean and Vince both have five successes and they both want the upper 80% confidence bound.

20. (a) The entire process must be one big set of Bernoulli trials. In particular, the \( p \) (not \( \hat{p} \)) must be the same for the past data and the future data. Also, the future data must be independent of the past data.

(b) First,

\[
r = 300/200 = 1.5 \text{ and } \hat{q} = 128/200 = 0.64.
\]

The 80% prediction interval is

\[
r \pm 1.282 \sqrt{r(1 + r) \hat{q}} = 1.5(72) \pm 1.282 \sqrt{1.5(72)(2.5)(0.64)} = 108 \pm 16.8.
\]