Examples of Applying Figure Sense in Statistics Problems

Example 1: In a large metropolitan area a study of the buying habits of typical customers led to the following observations. There are 52 percent female customers. 72 percent of the customers enjoy shopping for clothing. The percentage of females who enjoy shopping for clothing is 86 percent. What is the probability that a randomly chosen customer either enjoys shopping for clothing or is a female shopper?

Background:

Students are provided with a set of rules that can be useful in computing probabilities. For example these are the rules that are given in DSCI 202.

Probability Rules:

1) If all of the outcomes are equally likely, then the probability of an event, P(E), is equal to the number of events divided by the number of outcomes. <u>Events</u> are the specified results of a situation. <u>Outcomes</u> are all of the possible results of a situation.

2) Addition Law: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Read as P(A or B) = P(A) + P(B) – P(A and B)

3) Mutually exclusive events: If $P(A \cap B) = 0$, the events A and B are mutually exclusive events.

4) Complements: $P(A') = 1 - P(A)$

Read as P(not A) = 1 – P(A)

5) Conditional Probability: $P(B|A) = (P(A \cap B))/P(A)$

Read as P(B given A) = P(A and B)/P(A)

6) Multiplication Law: $P(A \cap B) = P(A)*P(B | A) = P(B)*P(A | B)$

If A and B are independent then $P(A \cap B) = P(A)*P(B)$

7) Events A and B are independent if and only if P(B |A) = P(B). Also P(A |B) = P(A).

8) Combinations: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

Read as n choose r.

9) Bayes Rule: $P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|B')P(B')}$

One issue in solving statistics word problems is to translate the work problem into symbols so that the student can recognize which probability rules to use to compute the answer. The following is an outline of how to use the figure sense habit of defining the problem to accomplish this transformation.

Figure Sense Habit: Define the Problem.

What do I know? or What information do I have to work with?

What do I want to accomplish?

What steps to I need to take to get from what I know to what I want to accomplish?

Step 1: What do I know?

List the events that can occur (assign a different symbol to denote each event).

F = the customer is female

S = the customer enjoys shopping

What do I know about these events? Notice that there are 3 numerical values in the problem, you should decide upon a symbol for each of these numerical values.

There are 52 percent female customers so $P(F) = .52$

72 percent of the customers enjoy shopping so $P(S) = .72$

The percent of females who enjoy shopping is 86 percent so $P(S \mid F) = .86$

Step 2: What do I want to find?

What is the probability that a randomly chosen customer either enjoys shopping or is a female?

In symbols this is: Find $P(S \cup F)$.

Step 3: What steps to I need to take to get from what I know to what I want to accomplish?

Using the addition law $P(S \cup F) = P(S) + P(F) - P(S \cap F)$.

We know $P(S)$ and $P(F)$ but we do not know $P(S \cap F)$.

Using the multiplication law $P(S \cap F) = P(F)*P(S \mid F)$.

We know $P(F)$ and $P(S \mid F)$.

Before solving the problem, we want to use another figure sense habit:

Figure Sense Habit: Look for unusual outcomes or exceptions.

Before solving the problem, ask: What do I expect the answer to be?

After solving the problem, ask: Is the answer consistent with what I expected?

If you find it difficult to determine the expected answer: After solving the problem ask: Does this answer make sense?

What do I expect the answer to be?

We want to find $P(S \cup F)$. We know $P(F) = .52$ and $P(S) = .72$.

Logically the probability of S or F occurring should be larger than the probability of S occurring or the probability of F occurring. Thus, $P(S \cup F)$ should be larger than .72. In addition, all probabilities must be less than 1.00. We expect $.72 \leq P(S \cup F) \leq 1.00$

Note: If it is difficult to figure out what you expect the answer to be before doing the problem, after you solve the problem you can ask: Is this answer reasonable?

Solve the problem:

$P(F) = .52 \quad P(S) = .72 \quad P(S \mid F) = .86$

$P(S \cap F) = P(F)*P(S \mid F) = (.52)*(.86) = .4472$

$P(S \cup F) = P(S) + P(F) - P(S \cap F). = .52 + .72 - .4472 = .7928$

Is this answer consistent with what I expected? Yes

Example 2: **An airline finds that 4 percent of the passengers that make reservations on a particular flight will not show up. Consequently, their policy is to sell 100 reserved seats on a plane that has only 98 seats. Find the probability that every person who shows up for the flight will find a seat available.**

Figure Sense Habit: Define the Problem.

What do I know? or What information do I have to work with?

What do I want to accomplish?

What steps to I need to take to get from what I know to what I want to accomplish?

Step 1: What do I know?

What is a random variable that is relevant for this problem? What symbol shall I use for this random variable?

X = the number of passengers that show up for the flight out of the 100 passengers who have reserved seats.

What do I know about the random variable X?

X has a binomial distribution with parameters n = 100 and $\pi = .96$ (the probability that each passenger shows up for the flight is .96).

Note for X to have a binomial distribution we have to make the assumption that the actions of all the passengers are independent each other. Is this assumptions reasonable?

This is a reasonable assumption if the passengers are not related to each other or if the passengers are not part of a group traveling together.

Step 2: What do I want to find?

Every person will find a seat if X is less than or equal to 98 so we want to find $P(X \leq 98)$.

Step 3: How do I find $P(X \leq 98)$?

The formula for computing probabilities for a Binomial random variable is: $P(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$

$P(X \leq 98) = 1 - \{P(X = 99) + P(X = 100)\}$

Figure Sense Habit: Look for unusual outcomes or exceptions.

Before solving the problem, ask: What do I expect the answer to be?

After solving the problem, ask: Is the answer consistent with what I expected?

What do I expect the answer to be?

If this is a realistic situation, I would expect the answer to be relatively large, since airlines would not like to deal with overbooked customers too often. I would expect $.90 \leq P(X \leq 98) \leq 1.00$.

Solve the problem:

$P(X = 99) = \binom{100}{99}(.96)^{99}(.04)^{1} = .0702$

$P(X = 100) = \binom{100}{100}(.96)^{100}(.04)^{2} = .0169$

$P(X \leq 98) = 1 - [\,.0702 + .0169] = .9129$

This is consistent with what I expected.

Example 3: **A bank want of provide a 99% daily service level at their ATM. This means that 99 days out of 100 days, the ATM will still have cash in it at the close of the day (10:00 pm). Of course, this also means that 1 day out of 100 days there will be customers who are turned away from the ATM due to insufficient cash for customer withdrawals. Past data illustrating the daily total of cash withdrawals is Normally distributed with mean daily total of $4125 and a standard deviation of $500. How much money must be in the ATM at the start of each day to achieve an approximate 99% service level?**

Figure Sense Habit: Define the Problem.

What do I know? or What information do I have to work with?

What do I want to accomplish?

What steps to I need to take to get from what I know to what I want to accomplish?

Step 1: What do I know?

What is the random variable relevant for this problem and what symbol will we use for this random variable?

Y = the daily cash withdrawals

What do I know about Y?

Y has a Normal distribution with mean 4125 and standard deviation 500.

Step 2: What do I want to find?

The amount of money k so that the probability of having enough cash is .99

Write this as a probability involving the random variable Y.

$P(Y < k) = .99$

Step 3: What steps to I need to take to get from what I know to what I want to accomplish?

Convert the problem about the random variable Y to a problem about a standard Normal Z.

$P(\frac{Y-\mu}{\sigma} < \frac{k-\mu}{\sigma}) = .99$

Find the value b in the Normal tables so that $P(Z < b) = .99$.

Set $b = \frac{k-\mu}{\sigma}$

Figure Sense Habit: Look for unusual outcomes or exceptions.

Before solving the problem ask: What do I expect the answer to be?

After solving the problem ask: Is the answer consistent with what I expected?

What do I expect the answer to be?

The amount of cash needs to be quite a bit above the mean withdrawn each day (4125). It should be more than 2 standard deviations above the mean so that would be larger than 4125 + 2*500 = 5125.

Solve the problem:

How do I find k so that P(Y < k) = .99

First we need to convert P(Y < k) to an expression about Z.

$P(Y < k) = P(\frac{Y-4125}{500} < \frac{k-4125}{500}) = P(Z < \frac{k-4125}{500}) = .99$

This means that $P(0 < Z < \frac{k-4125}{500}) = .49$

From the Normal tables P(0 < Z < 2.33) = .4901

Thus $\frac{k-4125}{500} = 2.33$. Solve this equation for k.

K = 4125 +(2.33)*(500) = 5290.

This is consistent with what was expected.

Example 4: **Recently, Fiat Motors has been advertising its new 5 year 50,000 mile warranty. The extended warranty covers engine, transmission, and drivetrain for all new Fiat-made cars for up to 5 years or 50,000 miles, which ever comes first. However one Fiat dealer believes the 5 year portion of the warranty is unnecessary since the mean number of miles driven by Fiat owners in 5 years exceeds 50,000 miles. A sample of 52 new Fiat owners had a sample mean of 51,117 miles driven with a sample standard deviation of 1,866 miles driven. Test whether the population miles driven is 50,000 or less so that both the 5 years and the 50,000 miles should be retained. Use α = .01.**

Figure Sense Habit: Define the Problem.

What do I know? or What information do I have to work with?

What do I want to accomplish?

What steps to I need to take to get from what I know to what I want to accomplish?

Step 1: What do I know?

What values are important in this problem and what symbols will be used for those values?

μ = the population number of miles driven in the first 5 years by Fiat owners.

We have a sample of n = 52 $\bar{x}$ (the sample mean) = 51117 s (the sample standard deviation) = 1866

Step 2: What do I want to find?

Is the population miles driven in the first 5 years less than or equal to 50,000?

What are the null and alternative hypotheses?

H_0: $\mu \leq 50{,}000$ (if this is correct they should maintain the current 5 years and the mileage)

H_1: $\mu > 50{,}000$

Step 3: What steps to I need to take to get from what I know to what I want to accomplish?

Compute the value of the test statistic: $t = \frac{\bar{x} - 50{,}000}{s/\sqrt{n}}$

Find the rejection region.

Determine if the test statistic is in the rejection region.

Figure Sense Habit: Look for unusual outcomes or exceptions.

Before solving the problem, ask: What do I expect the answer to be?

After solving the problem, ask: Is the answer consistent with what I expected?

What do I expect the answer to be?

Since 51,117 is quite a bit larger than 50,000 I expect to reject the null hypothesis.

Solve the problem: test these hypotheses?

Determine the test statistic: $t = \frac{\bar{x}-50{,}000}{s/\sqrt{n}}$

Determine the critical value: Since this is a one sided test, the critical value = $t_{\alpha,n-1} = t_{.01,51}$ = 2.403

Determine the rejection region: Reject H_0 if $t = \frac{\bar{x}-50{,}000}{s/\sqrt{n}} > 2.403$

Since $t = \frac{\bar{x}-50{,}000}{s/\sqrt{n}} = \frac{51{,}117-50{,}000}{1{,}866/\sqrt{52}} = \frac{1.117}{258.77} = 4.32 > 2.403$ reject H_0: $\mu \leq 50{,}000$.

This is consistent with what was expected.