Case 4 - Price Quotes:
Confidence Intervals and t-Tests for Paired Samples

Marlene Smith, University of Colorado Denver
Business School

Price Quotes:
Confidence Intervals and t-Tests for Paired Samples

Background

When an order is placed by a customer of a small manufacturing company, a price quote must be developed for that order. Because each order is unique, quotes must be established on an order-by-order basis by a pricing expert. The price quote process is labor-intensive, as prices depend on many factors such as the part number, customer, geographic location, market, and order volume. The sales department manager is concerned that the pricing process is too complex, and that there might be too much variability in the quoted prices. An improvement team is tasked with studying and improving the pricing process.

After interviewing experts to develop a better understanding of the current process, the team designed a study to determine if there is variability between pricing experts. That is, do different pricing experts provide different price quotes? Two randomly selected pricing experts, Mary and Barry, were asked to independently provide prices for twelve randomly selected orders. Each expert provided one price for each of the twelve orders.

The Task

Preliminary results of the study have been made available to you: Summary statistics of the price quotes from Mary and Barry (Exhibit 1):

Exhibit 1   Summary Statistics for Mary and Barry Price **Quotes**

Pricing Expert	Mean	Standard Deviation
Mary	$114.75	$11.05
Barry	$124.33	$20.70

Interpret these results for the manager in the context of the problem posed above. Because you have the data set, you can do any further analysis that you feel is warranted.
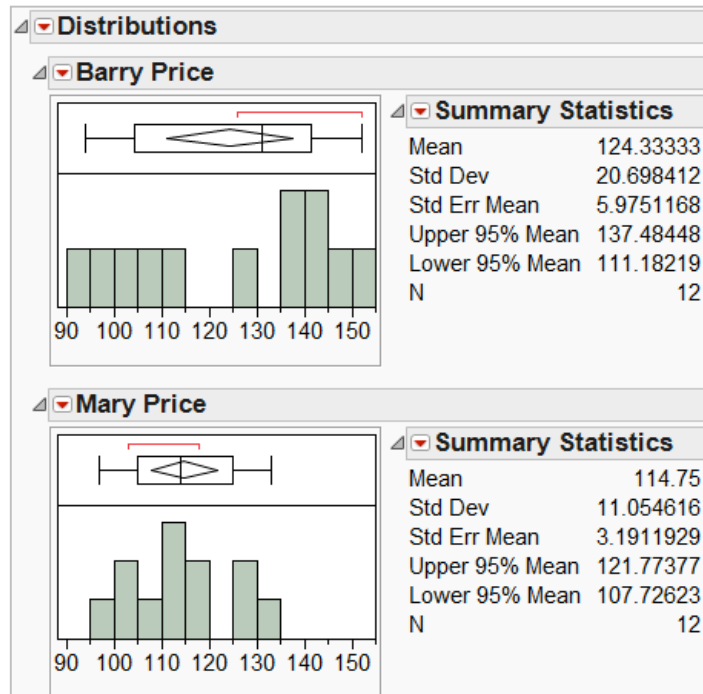
The Data   Price Quotes.jmp

The data set contains the order number, 1 through 12, and the price quotes by Mary and Barry for each order. Each row in the data set is the same order. Thus, Mary and Barry produced quotes for the same orders.

Analysis

A good first step might be to confirm the results provided in the Exhibit 1. Exhibit 2 shows that the means and standard deviations are indeed correct.

Exhibit 2   Summary Statistics of **Mary Price** and **Barry Price**

*(Analyze > Distribution; Use **Barry Price** and **Mary Price** as Y, Columns. Under the top red triangle, select Uniform Scaling.)*

The important question, though, is what to make of the summary statistics. Clearly, you would not want to make statements along the lines of, "Mary's quotes are too low," or "Barry's quotes are too high." The pricing study was not conducted to compare Mary's and Barry's quotes to some pre-determined "correct" price. We don't have a "gold standard" to compare them to.

Similarly, it would be incorrect to say, "Mary's quotes are lower than Barry's." Such a statement implies that Mary's quotes are *always* lower than Barry's. Mary's price quote for order #12, for example, was actually $11 *higher* than Barry's price. In fact, several of Mary's quotes were higher than Barry's.

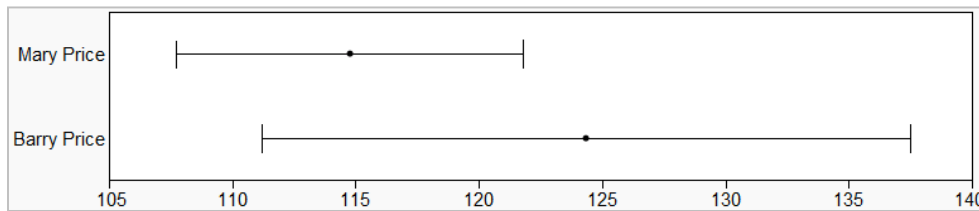Instead, we might want to determine whether Mary or Barry, *on average*, produce higher quotes.

A *numerically correct* interpretation of the averages reported in Exhibit 1 would be, "Mary's average quote is lower than Barry's average quote." But, comparing averages ignores sampling error. Thus, whether Mary's average quote is *statistically* lower than Barry's requires further investigation.

Statistically-trained professionals are aware that, had Mary and Barry been given twelve *different* orders to price, their averages would likely be different. Indeed, in a *different* sample of twelve orders, Mary's average price quote might have been higher, not lower, than Barry's. For that reason, we want to look not just at the averages but also their confidence intervals, which capture potential sampling error.

Confidence intervals are shown above in Exhibit 2 under "Moments" – Upper 95% Mean and Lower 95% Mean. The 95% confidence interval for Barry's average price quote is [$111, $137]; for Mary, [$108, 122].

Pictorially, we see that the 95% confidence intervals overlap (Exhibit 3): the numbers $111 through $122 are found in both intervals.

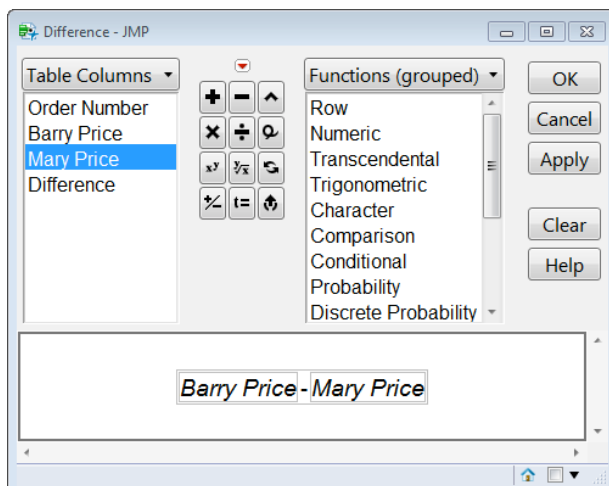**Exhibit 3**   Confidence Intervals for the Mean of **Mary Price** and **Barry Price**

*(Use Tables > Stack to stack **Mary Price** and **Barry Price**. Then, Graph > Graph Builder; drag and drop the **Data** column in X and the **Label** column in Y. Under Summary Statistics on the left select Mean, then under Error Bars select Confidence Interval. Alternatively, Graph > Chart can be used on the stacked data.)*

Confidence intervals take into consideration potential variation in averages from sample to sample. Here, the confidence intervals overlap. Thus, based on the 95% confidence coefficient, we don't have evidence to conclude that Barry's average price quote is statistically different from Mary's.

Now let's look at the standard deviations given in the preliminary report (and shown in Exhibit 2): $11.05 for Mary and $20.70 for Barry. What information do they provide? It is numerically correct to say that Mary has less variation in her quotes than Barry; i.e., that $11.05 is a smaller number than $20.70. However, these numbers aren't very meaningful *given the problem posed in the case*. For example, Mary's standard deviation tells us how much variation there is in Mary's quotes for the twelve orders. But, how did Barry price those same orders? In reality, we aren't concerned with how Mary's or Barry's quotes vary from order-to-order. What we really want to know is by how much Mary's quote for an order differs from Barry's quote for the same order.

Put another way, instead of looking at the individual price columns in the data set, we might compute a new variable, **Difference** (Exhibit 4), and compute a confidence interval for the average **Difference**.

Exhibit 4   Calculating **Difference** = **Barry Price** – **Mary Price**

*(Create a new column, and rename the column **Difference**. Right click on the column header, and select Formula to open the Formula Editor. To create the formula:*

*1. Select **Barry Price** from the columns list*
2. Select minus on the keypad
*3. Select **Mary Price** from the columns list*
4. Click OK.

To create this formula directly from the data table, select both columns in the data table, then right click and select New Formula Column > Combine > Difference.)
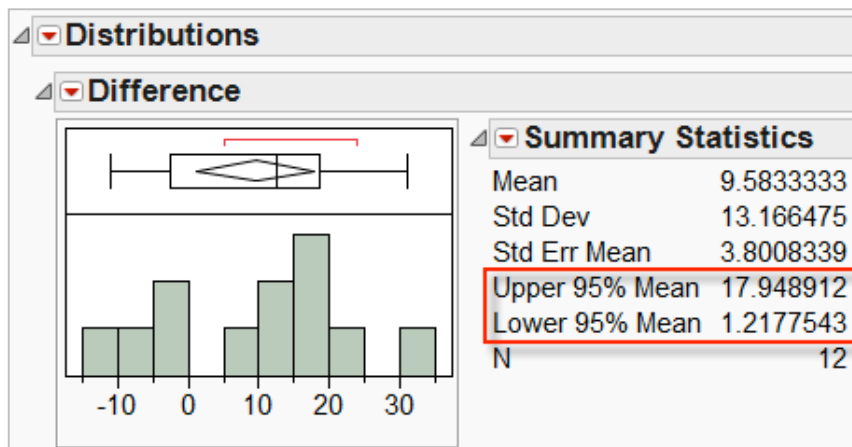
The resulting variable is shown in Exhibit 5 below. For Order Number 1, we see that Barry's price was $126 and Mary's was $114. Thus, the **Difference** is +$12.

Exhibit 5   Data set with **Difference**

Order Number	Barry Price	Mary Price	Difference
1	126	114	12
2	110	118	-8
3	138	114	24
4	142	111	31
5	146	129	17
6	136	119	17
7	94	97	-3
8	103	104	-1
9	140	127	13
10	152	133	19
11	108	103	5
12	97	108	-11

We can now compute the average **Difference**, and more importantly, a 95% confidence interval for the average **Difference** (Exhibit 6). Since the confidence interval, [$1.2, $17.9], does not contain zero, we can conclude that the average difference is statistically different from zero. In other words, there is a difference between prices quoted by Barry and Mary *for the same order* (when using a 95% confidence coefficient).
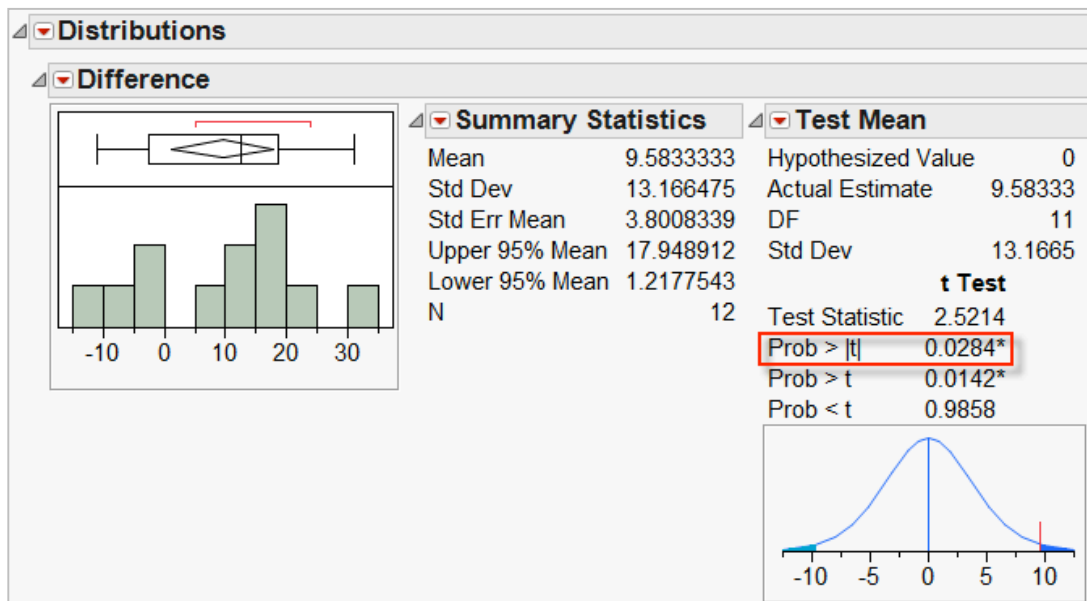
Exhibit 6   Confidence Interval for **Difference**

Summary Statistics

Mean	9.5833333
Std Dev	13.166475
Std Err Mean	3.8008339
Upper 95% Mean	17.948912
Lower 95% Mean	1.2177543
N	12

An alternative approach is to conduct a formal hypothesis test, a *One Sample t-Test*, to determine whether the average **Difference** is statistically different from zero (Exhibit 7). Since we are testing for differences between paired values, this test is special case of a One Sample t-Test known as a *Paired t-Test*. To help us make a decision, we use the first p-value provided: (Prob > |t|). If this p-value is less than 0.05, we can conclude that there is a statistically significant difference between the paired values. If this value is greater than 0.05, we cannot conclude that there is a difference. (Note that 0.05 is most commonly used as the cutoff for p-values. However, other values, such as 0.01 and 0.1 are sometimes used instead, depending on the context.)

In our case, we observe a p-value of 0.0284. This confirms what we learned with the confidence interval for the difference in Exhibit 6: There *is* a significant difference in average price quotes provided by Mary and Barry.

Exhibit 7   Hypothesis Test for **Difference**

*(Analyze > Distribution; Use **Difference** as Y, Columns and click OK. Click on the second red triangle, select Test Mean, and click OK. Or, for an equivalent test, use Analyze > Matched Pairs with **Barry Price** and **Mary Price** as Y, Paired Response.)*

Summary

Statistical Insights

Good data analysts recognize that all sample statistics are subject to sampling error; making managerial decisions based on summary statistics alone, without also accounting for sampling error, can lead to erroneous conclusions. Here, we might conclude that there is a problem with pricing consistency because Barry's average price quote is $10 higher than Mary's. But, could this difference be due to sampling error alone? Is this a *true* difference? Taking that extra leap, from numerical summaries to consideration of sampling error, is a key contribution of statistical thinking to workplace decision-making.

When testing for differences between matched pairs, as we saw in this case, a confidence interval for the average difference or a Paired t-Test is the way to go. Since both methods ignore order-to-order variation, they tend to be a more powerful for testing for paired differences, and have the added benefit of simplicity. In this case, both the confidence interval and the Paired t-Test led us to the conclusion that there is indeed a significant difference in quotes provided by Barry and Mary.

This study was designed to explore differences between pricing experts. Note that other sources of variation might also be present. For example, when providing a second quote for the same order (at some point later in time), does a pricing expert provide the same quote? An expanded study could be used to explore this additional source of variation.

Managerial Implications

After accounting for order-to-order variation, evidence showed that there is indeed variation in quotes provided by the pricing experts. A first step in improving the pricing process would be to explore these differences. Do the experts use different tools, software or calculations to prepare their quotes? Do they have more difficulty with some quotes than with others? Is additional training required? Are there opportunities to simplify and standardize the process? Are the pricing experts using best pricing practices?

A key point for managers is that the problem lies in the system rather than with the pricing experts. Only by studying and improving the system as a whole will the results improve.

JMP Features and Hints

This case uses the Distribution platform to produce summary statistics and confidence intervals. The Formula Editor was used to calculate a column of differences. We returned to Distribution to create a confidence interval for the average difference and to perform a One Sample t-Test. The Paired t-Test can also be performed from Analyze > Matched Pairs.

Early in the case we used Tables > Stack to stack the data. Stacking produced a new data table with the labels (the expert) in one column and the data (the price quotes) in a second column. We then used the Graph Builder on the stacked data to produce a visual display of the confidence intervals. Graph > Chart can also be used to create confidence intervals when the data is in this form. Note that, instead of using confidence intervals to compare the means for the two experts, we could use Fit Y by X to perform a Two Sample t-Test. This test allows us to formally test whether there is a significant difference between the means.

Exercises

1. The standard deviation of Mary's price quotes is $11.05. The standard error of the mean of Mary's price quotes is $3.19. Both are measures of variability.

    a. Carefully distinguish between these two numbers on the basis of how they are calculated and what they mean.
    b. Provide a nontechnical interpretation of each number.

2. Use the Medical Malpractice case and data set to determine whether there's a statistically significant difference between amounts paid for medical malpractice claims by gender.

    a. Use confidence intervals for averages, and compare the confidence intervals. Do the intervals overlap?
    b. As a bonus, use Fit Y by X to perform a Two Sample t-Test (the topic was not covered here, but give it a try anyway!).