

AP Statistics 2006: My responses. This is a first draft. Please comment, correct, suggest! Thanks, Bill Thill

1)

A) Both Catapult A and Catapult B have distances that are unimodal and symmetrically distributed. Catapult B shoots farther on average than Catapult A, because the center is higher (B's median = 138 cm vs A's median of 136 cm). Distances for catapult A are more variable than for B; the IQR for A = 4 but the IQR for B = 3). Furthermore, both the bottom and top 25% of distances for A are wider than the bottom and top 25% of distances for B.

B) Use catapult B. It has lower variability than catapult A. Therefore, the distances will be more reliably close to each other than they will for catapult A. They can set the catapult to the right distance, and 30 of 40 shots will be within 2.5 cm of the target: 10 of 40 shots are more than 2.5 cm away. This isn't true for A: 19 of 40 shots fall more than 2.5 cm from the center value.

C) Place it 138 cm from the target line. 138 is the median of all distances, so most of the time we'll be within 2.5cm above or below this value. Because the distribution is symmetric, this is the best setting, because it maximizes the chance that a shot is in the target area.

2)

A) "height" = height of suds, in mm. "grams soap" = amount of detergent in grams.
 $\widehat{height} = -2.679 + 9.500(\text{grams soap})$

B) The quantity s is the standard deviation of the residuals. On average, our predictions for soap height based on the linear model with amount of detergent are about 1.998 mm away from the actual height.

C) The SE of the slope = .7553. In this study, we estimate that each additional gram of detergent corresponds to about a 9.5 mm increase in suds. However, this estimate would vary upon repeated samples. The SE values tells us that in repeated samples, slope estimates like ours are, on average, about .7533 mm/gram off from the value of the slope of the "true" regression line between soap height and amt. of detergent for all possible runs of this study.

3.

a) *E is normally distributed. Therefore:*

$$P(M < 0) = P(E < -2). \quad P(E < -2) = P(Z < -1.333) \approx .0912$$

b) $P(\text{at least one is negative}) = 1 - P(\text{all three } M\text{'s are } \geq 0) = 1 - (1 - .0912)^3 = 0.2495.$

c) Variable: \bar{E} = sample mean of 3 errors. The average error value must be less than -2 for the mean measurement to be < 0. Therefore,

$$P(\bar{E} < -2) = P(Z < \frac{-2-0}{1.5/\sqrt{3}}) = P(Z < -2.31) \approx .01046$$

4.

We want to perform a **two sample t-interval** for the difference in mean wait times for ambulance – transported patients and self-transported patients.

Let:

μ_A = mean time, ambulances

μ_S = mean time, self – driven

99 % CI for $\mu_A - \mu_S$ requires that we have independent observations in both samples. We essentially have an SRS of both populations, so individual results are independent in both groups.

We have large sample sizes in each group, which allows us not to worry whether the populations are normally distributed. As long as there are no severe outliers in either population, the sampling distribution of $(\bar{X}_A - \bar{X}_B)$ will be roughly normal.

Our 99% CI for $\mu_A - \mu_S$ equals $(\bar{X}_A - \bar{X}_B) \pm t^*_{140.37} \sqrt{\frac{s_a^2}{n_a} + \frac{s_s^2}{n_s}} = (-4.291, -.2291)$

We are 99% confident that the difference in mean travel times for ambulance-driven heart attack patients is between 0.22 and 4.291 minutes faster than for self-driven heart attack patients.

b) Yes. Suppose there is no difference between mean wait times for the two groups. Then it would be very unlikely to get a 99% confidence interval which doesn't contain zero due to random chance alone: There's less than 1% chance of this happening. Therefore our results are unlikely to be due to chance, and therefore our results are statistically significant.

5.

a) 6 treatments:

Give nutrient A, tank has HIGH salinity.

Give nutrient A in a tank with LOW salinity.

Give nutrient B in a tank with HIGH Salinity.

Give nutrient B in a tank with LOW salinity.

Give nutrient C in a tank with HIGH salinity.

Give nutrient C in a tank with LOW salinity.

b) Number Shrimps 1-120. Arrange integers from 1-120 in random order (use a random number generator). The shrimps with those numbers in positions 1-10 of our list form treatment group #1. Take the next 10 #s off the list of random integers. The shrimps with those numbers for treatment group #2. In a similar fashion, go down the list in order until you have placed all the final 10 shrimp into group #12. Pick two integers from 1-12. The groups with those two numbers the two groups will receive nutrient A with HIGH salinity. Pick two more integers (without replacement) at random. The next two receive nutrient A with LOW salinity. Pick two more at random, without replacement. The next two receive nutrient B with high salinity. Pick again, the next two, B-Low. Pick again: the next two C-High, and the final two C-low. Weigh shrimps in each group before imposing treatments. Then place each group in one of the 12 tanks. This can be done in any way, assuming tanks are under similar conditions. Apply treatments for 3 weeks. After three weeks, measure weight again. Measure mean increase in weight for shrimp in each tank. Compare average weight gain among the six treatments.

c) By using only tiger shrimp, we will control variability in growth due to the breed of shrimp. For example, suppose I use "little popcorn" shrimp and "big jumbo" shrimp as experimental units as well. Popcorn shrimp will grow much less than tiger shrimp, and jumbo shrimp will grow much more than tiger shrimp, regardless of treatment. Even if these different breeds are equally distributed through all of the treatment groups, there will be so much variation in growth *within treatments* that we may not be able to detect whether differences between groups are statistically significant: there will be too much variation within groups that the differences in mean growth rates could easily be due to chance. Using tiger shrimps can be advantageous because by using only tiger shrimp, we reduce the amount of variability within treatment groups, making differences in mean results less likely to be due to chance.

d) By having only tiger shrimps in our samples, then any statistically significant results could be generalized only to the population of tiger shrimps. This is a disadvantage because our results don't apply as generally as we might like them to. This experiment would have to be replicated on different populations of shrimps (like popcorn shrimps, or jumbo shrimps) to determine if the conclusions could be generalized to other groups of shrimp.

#6:

a) Ho: $\sigma^2_x = 1.52$
 Ha: $\sigma^2_x > 1.52$

where σ^2_x represents the variance of the thermostat readings for recently manufactured thermostats.

b)
$$\frac{(n-1)s^2}{1.52} = \frac{(n-1) \left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)} \right)}{1.52} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{1.52} \approx 12.068$$

c) We trust our readings form a random sample of the population of all newly manufactured thermostats. If we can trust that that all other conditions for the test have been met, then we can use the fact that $P(\chi^2_9 > 12.068) \approx .20948$. Assuming that the variance hasn't changed, random chance would have produced a sample variance greater than 12.068 over 20% of the time. This isn't rare enough for us to reject the null. We fail to conclude that variance has gone up.

d) We need to find k such that $P(\chi^2_9 > k) = .05$. This critical value k is 16.92. The area shaded to the *right* of this value is the region covering all the values.

e) For each histogram, locate the bar where *test statistic* = 16.92. I shaded all bars higher than 15 on each histogram. Histogram III has the most area shaded, Histogram I has the second most shaded, and Histogram II has the smallest area shaded.

f) The population that has the largest variance will generate random samples that are *most* likely to generate a test statistic that would allow us to reject the null hypothesis and conclude that variance > 1.52 . For this reason, Sampling distribution III comes from a population with highest variance, because it has the most area shaded. Similarly, a population with the lowest variance will be *least* likely to generate a sample allowing us to reject the null. Therefore Histogram II illustrates the sampling distribution from a population with *least* variance, because the area shaded is smallest.

END