Molex Effect Size Problem

Current Scenario

- You oversee the Inside Sales group at Molex, in Lisle, Illinois.

- One of the key critical characteristics used to measure your group's performance is the cycle time (turnaround time) associated with the time required to respond to a 'Request for Bid (**RFB**)' to a customer.

Current Scenario

- When your group receives a RFB from an Original Equipment Manufacturer, OEM, it has a **maximum of 72 hours** to process the bid, conduct an analysis of the request, and return a bid.

- Any time this cannot be accomplished within 72 hours, Molex loses the opportunity to get the business associated with the product that is the subject of the RFB.

Current Scenario

- Over the last three years, Molex has been unable to bid on an average of $8.2 million worth of business each year, because the RFB process exceeded the 72 hour requirement.

- This is particularly troubling, because over those same three years, Molex was able to successfully acquire 60% of the business they bid on.

Current Scenario

- After a complete analysis of the RFB Cycle Time for the last three years, you are convinced that the current state of your capability (measured in hours) may be described as follows:
  - $\mu = 60$
  - $\sigma = 8$
  - $\gamma_3 = 0.0$
  - $\gamma_4 = 0.0$

Current Scenario

- Your personnel have convinced you to test a new hardware / software system which they believe will solve this Cycle Time problem.
- You are not so sure, but have agreed to conduct a short term experiment to find out.
- The **cost** of this new computer system is (one time) **$2 million.**

Current Scenario

- Management has made it clear that they do not wish to purchase this (or any other) system unless they can expect a minimal **ROI of 2:1**, after a **3 year period**.
- Further (as specifically related to this case), they want that ROI to be realized on the basis of business dollars acquired, not simply additional business bid on.

Current Scenario

- They have also made it very clear that they expect no more than a 5% probability that they will purchase this system, and subsequently discover that they are not getting the ROI expected.

1

Exercise 1

Effect Size Calculations for Means

Effect Size for Means

Assumptions:

- The System (Solution) is intended to affect the Average Cycle Time
- Variability and shape will be unaffected (holding them constant for simplicity)
- The system will either have a beneficial effect, or no effect at all (directional effect)

Effect Size for Means

Goal:

- Determine the Effect Size required for this situation.
- Be sure you identify the items A through G from this presentation.

H_0: $\mu \geq 60$
H_1: $\mu < 60$

Effect Size for Means

- **A** is the total annual loss (estimated loss of business) \$8,200,000.00 * 0.6 (60%) = \$4,920,000.00 annual loss.

- **B** is the cost of the solution is \$2,000,000.00, a 2 million dollar computer system .

- **C** is the Solution Benefit = solution cost with the ROI requirement applied (\$2,000,000.00 * 2) / 3 = \$1,333,333.33 per year (ROI for this case is a 2 to 1 in 3 years pay back.)

Effect Size for Means

- **D** is the Percent of time they fail to meet the 72 hour maximum time to provide the quote. (% area of normal curve with $\mu = 60$ and $\sigma = 8$ that is above 72) = 6.6807% [use 4 decimal places in these calculations]

- **E** is the Dollar loss per 1% out of specification events. A/D = $736,447.56/1%

Effect Size for Means

- **F** is the Maximum allowable annual loss they are willing to tolerate (given that the solution with ROI is on the way.
  A-C = $3,586,666.67

- **G** is the New maximum allowable defect rate (obtained from dollar max and $/1% out):
  F/E = 4.8702%

Effect Size for Means

- Then, you have to take this percentage from **G** and from the normal distribution obtain a z-score.

$$Z = \frac{X - \mu}{\sigma}$$

- Make sure you identify the side of the distribution (`lower.tail = T or lower.tail = F`) you are on so you can get the correct z-score. The plus or minus will make a difference.)

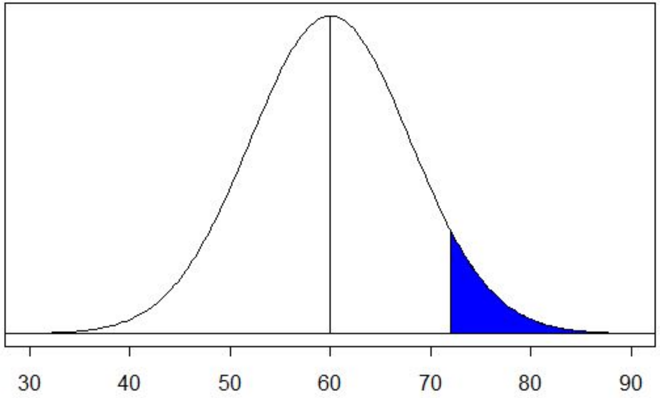
Effect Size for Means

- This z-score will be needed for calculating both the effect for the mean, and for calculating the effect for the standard deviation.

- The effect size for proportions, uses only the initial defect rate (current), and the maximum allowable defect rate in order for the solution benefit (cost with ROI applied) to be effective (new).

Current

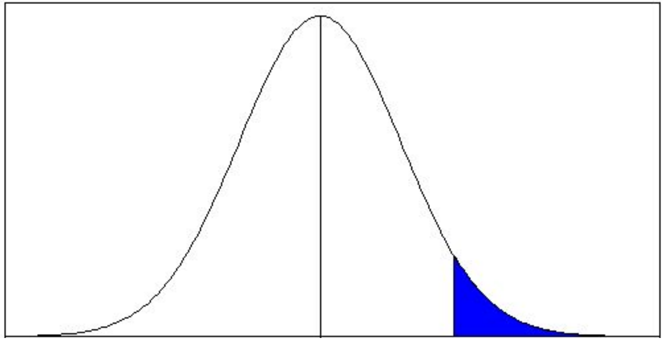
1% = $736,447.56



Loss = $4,920,000
(60% of $8,200,000)

Mean = 60
Std Dev = 8
USL = 72
Z-Score = 1.5
% above USL = 6.6807%

New



Loss = $3,586,666.67

Area = 4.8702%

Mean = 0
Std Dev = 1
Area Above = 4.8702%
Z-Score = 1.6576

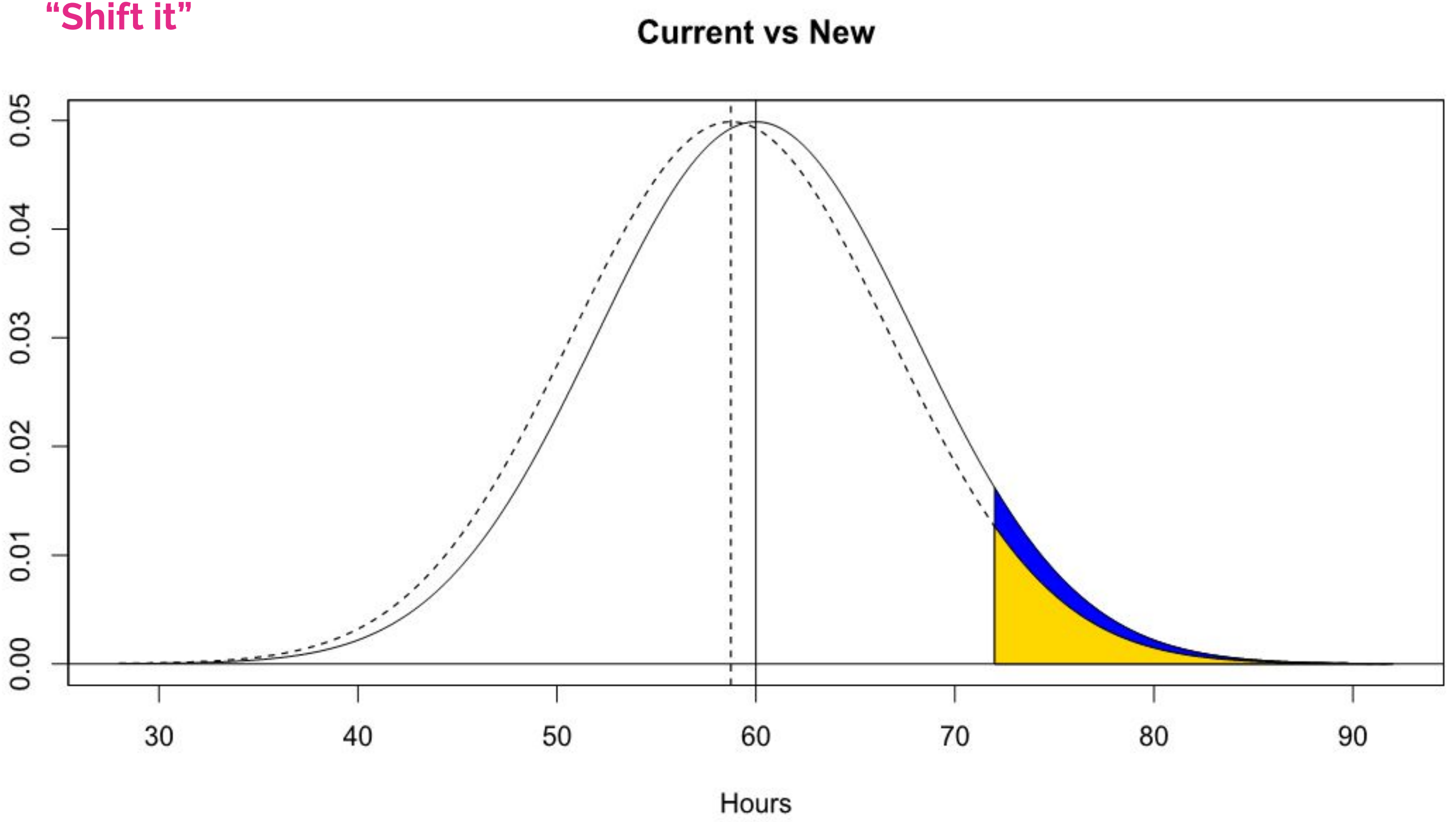
Effect Size = $\Delta\mu$ = -1.2606

$1.6576 = (72-\mu_{new})/8$

$\mu_{new} = 58.7395$

$\Delta\mu = \mu_{new} - \mu_{current}$

"Shift it"



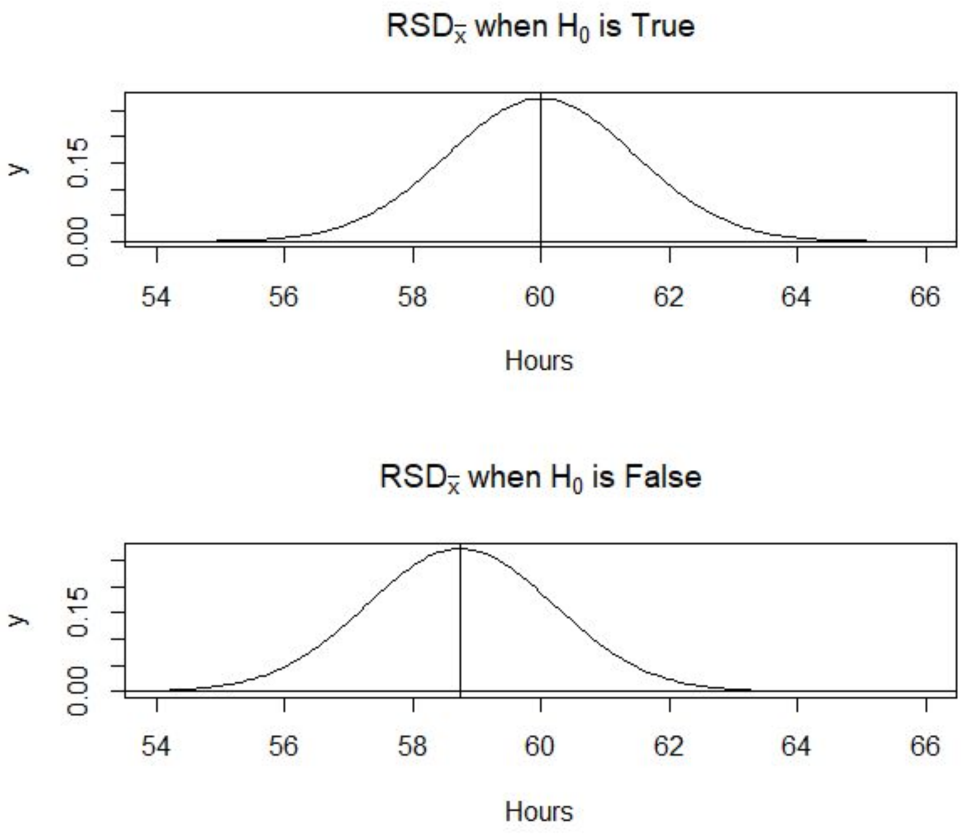
Effect Size for Means

- Assuming that you have been allowed to select a random sample of 30 incoming RFBs for processing through the proposed Hardware / Software System, what is the power of this test?

Effect Size for Means

- Note, now we have to think about the **RSD of means** for samples of size **n=30**.

- The good news is that lolcat will take all of that into account when you use the Sample Size routines to compute Power and Sample Size as demonstrated in class.



H_0: $\mu \geq 60$
H_1: $\mu < 60$

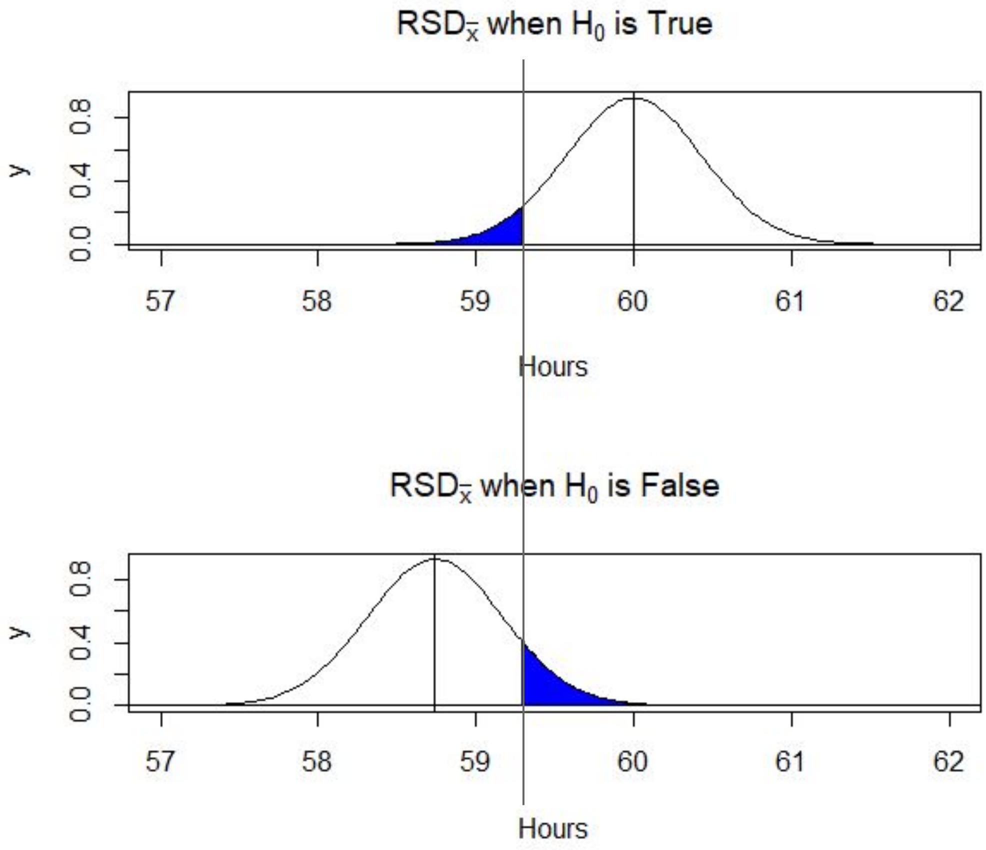
```
nqtr(power.mean.t.onesample(
sample.size = 30,
effect.size = -1.2606,
variance.est = 8^2,
alpha = 0.05,
alternative = "less"),4)

test         t
type         one.sample
alternative  less
sample.size  30
actual       30
df           29
effect.size  -1.2606
variance     64
alpha        0.05
conf.level   0.95
beta         0.7887
power        0.2113
```

Effect Size for Means

- Assuming that management wants Power to be a minimum of 90%, what is the minimum sample size you would use?

- See the following slide for details, and always remember to use the “t” version for power and n calculations in these situations.




```
nqtr(sample.size.mean.t.onesample(
effect.size = deltamu
,variance.est = sd^2
,alpha = alpha
,beta = beta
,alternative = "less"),4)
```

```
test        t
type        one.sample
alternative less
sample.size 347
actual      347
df          346
effect.size -1.2606
variance    64
alpha       0.05
conf.level  0.95
beta        0.1
power       0.9005
```

Effect Size for Means

- What if we needed to run a **two–tailed** test instead?
- What is the hypothesis we are testing?
  H_0: $\mu = 60$
  H_1: $\mu \neq 60$
- The answer to this question depends on whether Type I error was originally set based strictly on the probability of committing the error in a single direction; i.e. avoiding ONLY the error associated with thinking that was less than 60, when it was not.

Effect Size for Means

- If this is the case, and if when offered a two-tailed test option, management insisted on maintaining this level of protection (a less conservative option), then we would double the magnitude of the original Type I error level for the two-tailed test calculation.

Effect Size for Means

```
nqtr(sample.size.mean.t.onesample(effect.size = deltamu, variance.est = sd^2, alpha =
alpha*2, beta = beta, alternative = "two.sided"),4)
```

```
test        t
type        one.sample
alternative two.sided
sample.size 347
actual      347
df          346
effect.size -1.2606
variance    64
alpha       0.1
conf.level  0.9
beta        0.1
power       0.9005
```

Of course, the doubling of Type I Error accompanied by moving from a One-Tailed to Two-Tailed test results in no change to the sample size required, versus a one-tailed test with half the **α**.

Just remember this rule: Alpha is alpha is alpha overall, and then properly determine whether you have directional or non-directional hypotheses (or situation).

Based upon the strict interpretation of the original problem of the statement, this would be the correct answer to this question (which is twice the alpha for a one tailed test).

Effect Size for Means

- On the other hand, if the origin of the original was associated with an error in either direction and both were possible (of course, if both were possible we would not have run a one-tailed test to begin with)

- OR if when faced with a total of a 10% Type I Error Level and only 90% Confidence, management were to 'back off' of the original requirement that the 5% error level on one side only be maintained, (which is the same as 10% in two tails), then the original would be maintained, but split into two equal rejection regions:

Effect Size for Means

```
nqtr(sample.size.mean.t.onesample(effect.size = deltamu, variance.est = sd^2, alpha =
alpha, beta = beta, alternative = "two.sided"),4)
```

```
test        t
type        one.sample
alternative two.sided
sample.size 426
actual      426
df          425
effect.size -1.2606
variance    64
alpha       0.05
conf.level  0.95
beta        0.1
power       0.9006
```

So, if = 5% overall, and you are assessing a non-directional hypothesis you would select a “two-tail” test as just shown. Of course, this would also be the calculated sample size if the original statement of the problem had been posed as (for example): ‘Management wishes to restrict the probability of incorrectly inferring that “a change of any type” has taken place at a maximum of 5%’.

2

Exercise 2

Effect Size Calculations for Variability

Effect Size for Std Dev

Assumptions:

- The System (Solution) is intended to affect the variability of cycle time
- The mean and shape of the distribution will be unaffected; and
- The System will either have a beneficial effect, or no effect at all

Effect Size for Std Dev

- We can see that the exact same information is available and applies with respect to the items A, B, C, D, E, F and G.
- And, this means that the computed z-score of 1.6576 applies as well.

$$Z = \frac{X - \mu}{\sigma}$$

Effect Size for Std Dev

- Now, however, we need to use it (the z-score) to solve for a new Standard Deviation, new by use of the following formula:

  $Z = (SL - \mu)/\sigma_{new}$

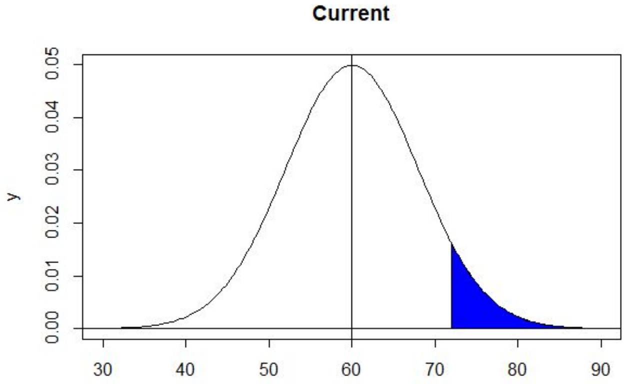
  $1.6576 = (72 - 60)/\sigma_{new}$

  $\sigma_{new} = (72-60)/1.6576 = 7.2394$

$$Z = \frac{X - \mu}{\sigma}$$

Current

1% = $736,447.56



Loss = $4,920,000

Area = 6.6807%

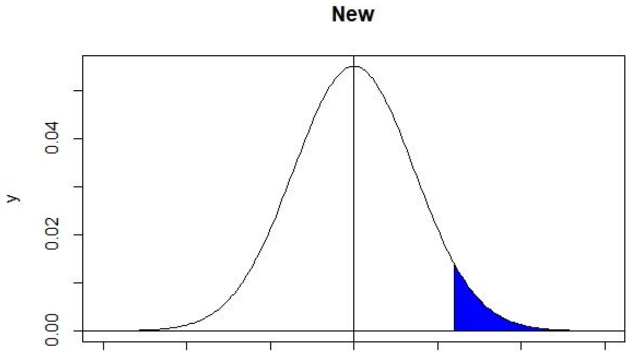
New

Effect Size = $\Delta\sigma$ = 0.7605

$1.6576 = (72-60)/\sigma_{new}$

$\sigma_{new} = 7.2395$

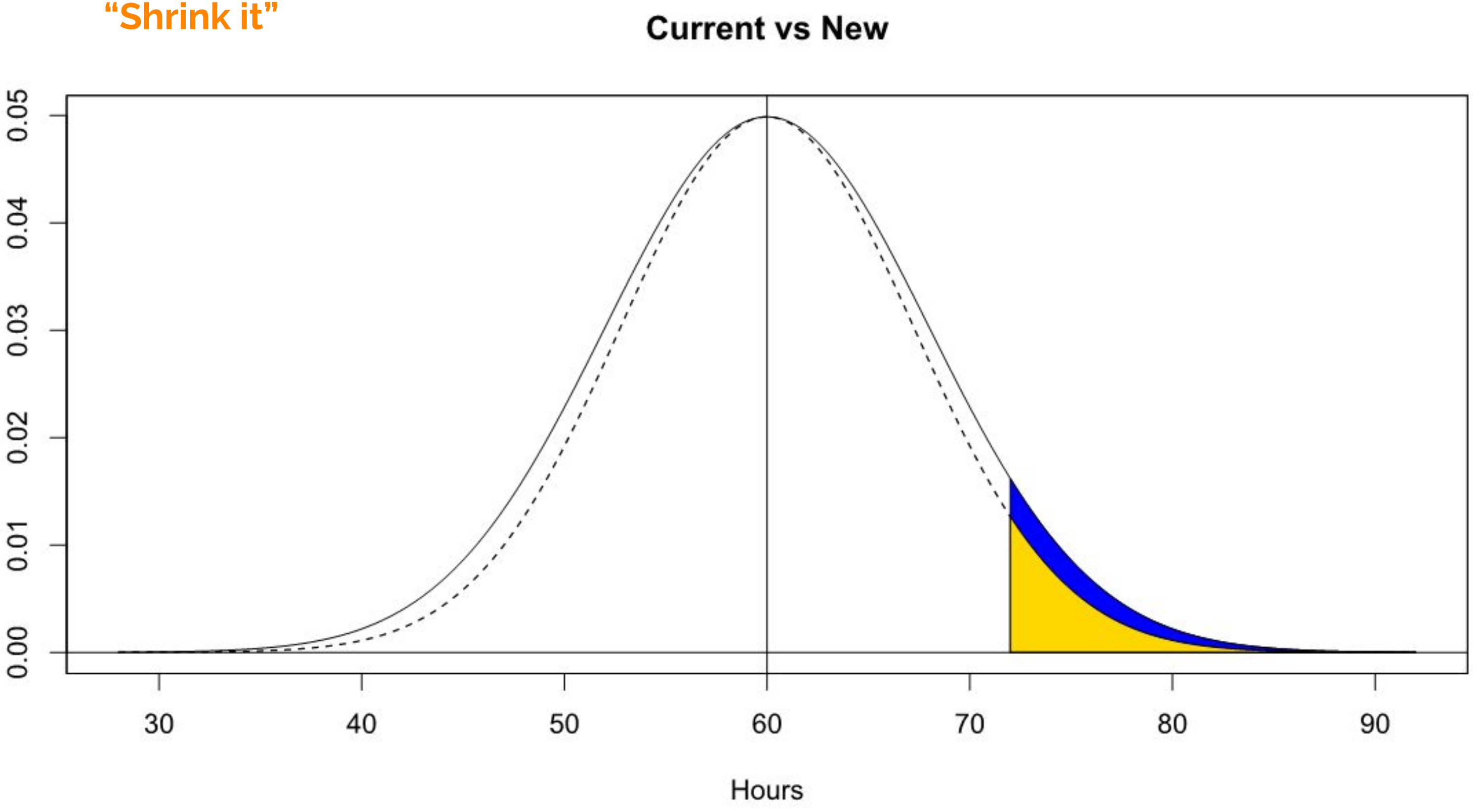
$\Delta\sigma = \sigma_{current} - \sigma_{new}$



Loss = $3,586,666.67

Area = 4.8702%

"Shrink it"



Effect Size for Std Dev

Now, with the $\Delta\sigma$, and assuming that:

- Management wants no more than a 5% chance of purchasing the proposed system, and subsequently discovering that it is not returning the required ROI (Type I Error); and

- Management wants no more than a 10% chance of having the proposed system incorrectly rejected due to sampling error (a Type II Error), then:

Effect Size for Std Dev

- What is the minimum sample size required for a one-tailed test?
- What are the Hypotheses we would be testing?

$H_o : \sigma^2 \geq 64$

$H_1 : \sigma^2 < 64$

- Note, we would only reject Ho when if we obtained a new standard deviation which is smaller than 8 by more than the effect size which would be the original Standard Deviation minus the effect size!)

Effect Size for Std Dev

```
nqtr(sample.size.variance.onesample(null.hypothesis.variance = sd^2
,alternative.hypothesis.variance = sdnew^2
,alpha = alpha
,beta = beta
,alternative = "less"),4)
```

```
test        chi-square
type        one.sample
alternative less
sample.size 435
df          434
ratio       0.8189
alpha       0.05
conf.level  0.95
beta        0.0997
power       0.9003
```

Effect Size for Std Dev

- What is the minimum sample size required for a **two-tailed** test?
- What are the Hypotheses we would be testing?

H_0: $\sigma^2 = 64$

H_1: $\sigma^2 \neq 64$

Effect Size for Std Dev

- Assuming that the position established by management in the original statement of the problem is maintained (that is, that a maximum of a 5% risk of inferring the change 'worked' when, in reality, it did not change), then the sample size (just as with the means) remains unchanged from the calculation associated with the one-tailed test:

Effect Size for Std Dev

```
nqtr(sample.size.variance.onesample(null.hypothesis.variance = sd^2
,alternative.hypothesis.variance = sdnew^2
,alpha = alpha*2
,beta = beta
,alternative = "two.sided"),4)
```

```
test        chi-square
type        one.sample
alternative two.sided
sample.size 435
df          434
ratio       0.8189
alpha       0.1
conf.level  0.9
beta        0.0997
power       0.9003
```

This would be the correct answer **if** <u>the Type I Error level was restricted to 5% for a *single-sided mistake*,</u> as specified in the original statement of the problem.

Effect Size for Std Dev

- If the change in the variance occurred in the opposite direction (now that we are running a two-tailed test, we'd have to consider the possibility) under the same assumption (i.e. holding Type I error at a maximum of 5% in either direction as opposed to 5% total or overall, we would have:

Effect Size for Std Dev

```
nqtr(sample.size.variance.onesample(null.hypothesis.variance = sd^2
,alternative.hypothesis.variance = (sd + deltasd)^2
,alpha = alpha*2
,beta = beta
,alternative = "two.sided"),4)
```

```
test         chi-square
type         one.sample
alternative  two.sided
sample.size  518
df           517
ratio        1.1992
alpha        0.1
conf.level   0.9
beta         0.0996
power        0.9004
```

Notice: Here, the `sd^2` is the same as before, that is the current, and the `(sd + deltasd)^2` is a "new variance" resulting from a "larger or + effect size. So, $\sigma_{current} + \Delta\sigma = 8 + 0.7606 = 8.7606$ resulting in a variance ratio of 1.1992 which is larger than 1.

Effect Size for Std Dev

- On the other hand, if Type I Error was maximized at an overall total of 5%, then running this analysis as a two-tailed test (if the standard deviation were to ***decrease*** by the effect size of 0.7606) would yield:

Effect Size for Std Dev

```
nqtr(sample.size.variance.onesample(null.hypothesis.variance = sd^2
,alternative.hypothesis.variance = sdnew^2
,alpha = alpha
,beta = beta
,alternative = "two.sided"),4)
```

```
test        chi-square
type        one.sample
alternative two.sided
sample.size 536
df          535
ratio       0.8189
alpha       0.05
conf.level  0.95
beta        0.0999
power       0.9001
```

Effect Size for Std Dev

- However, if the standard deviation were to ***increase*** by the effect size of 0.7606) would yield:

Effect Size for Std Dev

```
nqtr(sample.size.variance.onesample(null.hypothesis.variance = sd^2
,alternative.hypothesis.variance = (sd + deltasd)^2
,alpha = alpha
,beta = beta
,alternative = "two.sided"),4)
```

```
test        chi-square
type        one.sample
alternative two.sided
sample.size 631
df          630
ratio       1.1992
alpha       0.05
conf.level  0.95
beta        0.1
power       0.9
```

Notice: The difference in the ratio for the two results has to do with the numerator term
being either current + or – the effect size, $\Delta\sigma$

3

Exercise 3

Effect Size Calculations for Proportions

Effect Size for Proportions

- The vendor wanting to sell us the new system is willing to tell us only whether a randomly sampled RFB was or was not returned within 72 hours (Binary data related to Proportions). If this is the way the experiment must be conducted, and

Effect Size for Proportions

- Management wants no more than a 5% chance of purchasing the proposed system, and subsequently discovering that it is not returning the required ROI (although we will also illustrate the two-tailed situation where management wants no more than a 5% probability of an incorrect rejection of a true null hypothesis) ; and

Effect Size for Proportions

- Management wants no more than a 10% chance of having the proposed system incorrectly rejected due to sampling error; then:

- What is the minimum sample size required for a one-tailed test?

Effect Size for Proportions

- In this case, associated with proportions, it is easy because we just work directly with the initial or current proportion out of specification and the new proportion required in order to make the payout acceptable.

- So, we only need current and new, to obtain the effect size we seek.

Effect Size for Proportions

- Once we have it, we can use lolcat to obtain our sample size and power values as appropriate.
- But first, let's make sure we have calculated our effect size for proportions, $\Delta\pi$.
- Per our previous definitions, $\Delta\pi$ will be equal to the absolute value of the $\pi_{current} - \pi_{new}$, so $\Delta\pi = 0.066807 - 0.048702 = 0.018105$.

Effect Size for Proportions

- Thus, for determining the sample size for detecting a ***reduction*** in defect rate we would use current for our null.hypothesis.proportion and $\pi_{current} - \Delta\pi$ for alternative.hypothesis.proportion in lolcat, and that would give us the following:

Effect Size for Proportions

```
nqtr(sample.size.proportion.test.onesample.exact(null.hypothesis.proportion = p0
,alternative.hypothesis.proportion = p1
,alpha = alpha
,beta = beta
,alternative = "less"),4)
```

```
test                              proportion
type                              one.sample
alternative                       less
sample.size                       1434
actual                            1434
null.hypothesis.proportion        0.0668
alternative.hypothesis.proportion 0.0487
alpha                             0.05
conf.level                        0.95
beta                              0.0974
power                             0.9026
```

Effect Size for Proportions

- What are the Hypotheses we would be testing?

  H_0: $\pi \geq 0.066807$
  H_1: $\pi < 0.066807$

Effect Size for Proportions

- What is the minimum sample size required for a **two-tailed** test, if the risk of buying a new process that doesn't work remains at a maximum of 5% for either side (therefore requiring a total Type I error level of 10%)?

- What are the Hypotheses we would be testing?

  $H_0: \pi = 0.066807$

  $H_1: \pi \neq 0.066807$

Effect Size for Proportions

```
nqtr(sample.size.proportion.test.onesample.exact(null.hypothesis.proportion = p0
,alternative.hypothesis.proportion = p1
,alpha = alpha*2
,beta = beta
,alternative = "two.sided"),4)
```

```
test                              proportion
type                              one.sample
alternative                       two.sided
sample.size                       1434
actual                            1434
null.hypothesis.proportion        0.0668
alternative.hypothesis.proportion 0.0487
alpha                             0.1
conf.level                        0.9
beta                              0.0974
power                             0.9026
```

Effect Size for Proportions

```
p2<-p0 + deltap
nqtr(sample.size.proportion.test.onesample.exact(null.hypothesis.proportion = p0
,alternative.hypothesis.proportion = p2
,alpha = alpha*2
,beta = beta
,alternative = "two.sided"),4)
```

```
test                              proportion
type                              one.sample
alternative                       two.sided
sample.size                       1822
actual                            1822
null.hypothesis.proportion        0.0668
alternative.hypothesis.proportion 0.0849
alpha                             0.1
conf.level                        0.9
beta                              0.0992
power                             0.9008
```

Based on the specific statement in the original statement of the problem, these calculations would represent the correct answers. (Note, you'd always pick the larger sample size just in case!)

Effect Size for Proportions

- If, on the other hand, additional discussions with management led to a decision to maximize the total Type I Error level at 5%, then we would have the following:

Effect Size for Proportions

```
nqtr(sample.size.proportion.test.onesample.exact(null.hypothesis.proportion = p0
,alternative.hypothesis.proportion = p1
,alpha = alpha
,beta = beta
,alternative = "two.sided"),4)
```

```
test                               proportion
type                               one.sample
alternative                        two.sided
sample.size                        1761
actual                             1761
null.hypothesis.proportion         0.0668
alternative.hypothesis.proportion  0.0487
alpha                              0.05
conf.level                         0.95
beta                               0.0987
power                              0.9013
```

Effect Size for Proportions

```
p2<-p0 + deltap
nqtr(sample.size.proportion.test.onesample.exact(null.hypothesis.proportion = p0
,alternative.hypothesis.proportion = p2
,alpha = alpha
,beta = beta
,alternative = "two.sided"),4)
```

```
test                              proportion
type                              one.sample
alternative                       two.sided
sample.size                       2230
actual                            2230
null.hypothesis.proportion        0.0668
alternative.hypothesis.proportion 0.0849
alpha                             0.05
conf.level                        0.95
beta                              0.0989
power                             0.9011
```

(Note, in this case we would have to do two calculations, one for subtracting the effect size, $\Delta\pi$, from $\pi_{current}$ to get p1, and then adding $\Delta\pi$ to $\pi_{current}$ to get p2 for the second one. These lead to the two results that are provided on this slide and the prior slide.)