Name_____ **Date**_____ **Class**_____

WISE Power Tutorial – Answer Sheet

Power: The B.E.A.N. Mnemonic

Select true or false for each scenario:

(Assuming no other changes)	True	False
1. As effect size increases, power decreases.	☐	☐
2. As sample size increases, power increases.	☐	☐
3. As alpha error increases, power decreases.	☐	☐
4. Beta error is unrelated to power.	☐	☐

Exercise 1a: Power and Mean Differences (Large Effect)

To simulate drawing one sample of 25 cases, press **Sample**. The mean and *z*-score are shown in the applet (bottom right box). Record these values in the first pair of boxes below (you may round the mean to a whole number).

Trial	1	2	3	4	5	6	7	8	9	10
Mean				579	574	594	600	541	585	578
Z-Score				3.96	3.72	4.69	4.99	2.04	4.23	3.92

1a. How many times could you reject the null hypothesis in your ten samples?
(With one-tailed alpha $\alpha = .05$, $z = 1.645$, so reject H_0 if your *z*-score is greater than 1.645)

Exercise 1b: Power and Mean Differences (Small Effect)

I predict that statistical power for the test of the DEUCE program compared to the test of the ACE program will be:

☐ Less ☐ The Same ☐ Greater

To simulate drawing a sample of 25 from graduates from the DEUCE program, enter the following information into the WISE Power Applet:

- μ_0 = 500 (**null mean**);
- μ_1 = 520 **(alternative mean)**;
- σ = 100 (**standard deviation**);
- α = .05 (**alpha error rate, one tailed**);
- **n** = 25 (**sample size**).

Do three simulations of drawing a sample of 25 cases, and record the results below.

Trial	1	2	3	4	5	6	7	8	9	10
Mean				509	511	513	502	492	513	533
Z-Score				0.44	0.54	2.06	0.11	-0.41	0.65	1.63

1c. What is the power for this test as shown in the applet? ☐

1d. How many of your ten simulated samples allowed you to reject the null hypothesis? (Use one-tailed alpha α = .05, z = 1.645, so reject H_0 if your z-score is greater than 1.645)
☐

1e. For the ACE program, the effect size was **.8** and the power of the statistical test was **.991**; what can you conclude about the relationship between effect size and power?

☐ **A.** The test for the ACE program, which had a larger effect size, had more power.

☐ **B.** The test for the DEUCE program, which had a smaller effect size, had more power.

☐ **C.** Effect size is unrelated to power.

Exercise 1c: Power and Variability (Standard Deviation)

1f.I think that with a smaller standard deviation in the population, the statistical power will be

☐ Less ☐ Unchanged ☐ Greater ☐ I don't know

To simulate drawing a sample from graduates of the TREY program that has the same population mean as the DEUCE program (520), but a smaller standard deviation (50 instead of 100), enter the following values into the WISE Power Applet:

- μ_0 = 500 (**null mean**);
- μ_1 = 520 **(alternative mean**);
- σ = 50 (**standard deviation**);
- α = .05 (**alpha error rate, one tailed**);
- **n** = 25 (**sample size**).

Do three simulations of drawing a sample of 25 cases and record the results below.

Trial	1	2	3	4	5	6	7	8	9	10
Mean				512.1	516.4	515.6	515.4	525.2	535.3	528.6
Z-Score				1.21	1.64	1.56	1.36	2.52	3.53	2.86

1g. How many of your ten simulated samples allowed you to reject the null hypothesis? (Use one-tailed alpha α = .05, z = 1.645, so reject H_0 if your z-score is greater than 1.645)

1h. What is the power for this test (from the applet)?

1i. In Exercise 1b the DEUCE program had a mean of 520 just like the TREY program, but with samples of N = 25 for both programs, the test for the DEUCE program had a power of .260 rather than .639. The standard deviation for DEUCE was 100 rather than 50. Why is statistical power greater for the TREY program?

☐ Because smaller population variance always produces greater power.

☐ Because the program with the larger effect size always produces greater power.

☐ Neither of these reasons is sufficient.

Exercise 1d: Summary of Power and Effect Size

Below are key statistics for each of two new training programs, SLAM and DUNK.

Statistics for SLAM	Statistics for DUNK
μ_0 = 500 **(null mean)**; μ_1 = 540 **(alternative mean)**; σ = 50 **(standard deviation)**; α = .05 **(alpha error, one tailed)** n = 50 **(sample size)**.	μ_0 = 500 **(null mean)**; μ_1 = 520 **(alternative mean)**; σ = 20 **(standard deviation)**; α = .05 **(alpha error, one tailed)** n = 50 **(sample size)**.

Do you expect that a test of statistical significance would have greater power for the SLAM program or the DUNK program? Why? Respond to the following true/false statements. You can hover over the "Check" boxes below to check your answer. See if you can answer all statements correctly before you check your answers.

T	F	
☐	☐	**1j.** The test of the SLAM program will have greater power because SLAM has a larger mean than DUNK.
☐	☐	**1k.** The test of the SLAM program will have greater power because SLAM has a larger standard deviation than DUNK.
☐	☐	**1l.** The test of the DUNK program will have greater power, because a program with smaller standard deviation has greater power.
☐	☐	**1m.** The test of the SLAM program will have greater power because SLAM has a greater effect size.
☐	☐	**1n.** The test of the DUNK program will have greater power than the test for the SLAM program if sample size and alpha are the same, because DUNK has a greater effect size.
☐	☐	**1o.** Power is the same for the tests of the two programs because the samples have the same size.
☐	☐	**1p.** Holding all other factors constant, a larger difference between the null and alternate population means will always yield greater power.
☐	☐	**1q.**Holding all other factors constant, power is greater when the variance of the null and alternate populations is greater.

Exercise 2: Power and Sample Size

In Exercise 1d we drew samples of 25 graduates from the DEUCE program but in Exercise 2 we will draw samples of n = 100 and n = 4. Watch what happens and think about how you can explain how statistical power is influenced by sample size.

For this exercise, do ten simulations of drawing a sample of 100 cases and record the results below. You don't need to record means; just select the button for "Reject H_0" or "Fail to Reject" for each of the ten simulations.

	1	2	3	4	5	6	7	8	9	10
Reject H_0	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fail to Reject	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

2a. Power (as shown in the applet) = ☐

Next, do ten simulations of drawing a sample of 4 cases and record the results below.

(Set sample size n = 4, press Enter)

	1	2	3	4	5	6	7	8	9	10
Reject H_0	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fail to Reject	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

2b. Power (as shown in the applet) = ☐

2c. How many times out of 10 did you Reject H_0 for each of the two scenarios?

n = 4: ☐
n = 100: ☐

If nothing else is changed, power is greater with a larger sample size because:

(select True or False for each before you check your answers):

T F

- ☐ ☐ 2d. The effect size is larger.
- ☐ ☐ 2e. The alpha error is smaller.
- ☐ ☐ 2f. The alpha error is larger.
- ☐ ☐ 2g. The population variance is smaller.
- ☐ ☐ 2h. The variance of the alternate sampling distribution is smaller.
- ☐ ☐ 2i. More of the alternate sampling distribution exceeds the critical value.

2j. Explain to a classmate why statistical power increases as the sample size increases.

Exercise 3: Power and Alpha

For this example, use one-tailed alpha $\alpha = .01$ ($z = 2.326$). In this case, we will reject the null hypothesis only if a sample mean is so large that it would occur less than 1% of the time given the null hypothesis is true. You do not need to draw additional samples for this problem; you can use the data recorded for samples drawn in **Exercise 1** ($\mu_0 = 500$, $\sigma = 100$, $n = 25$, $\alpha = .05$, $z = 1.645$).

Data from Exercise 1

ACE Program ($\mu_1 = 580$)

Trial	1	2	3	4	5	6	7	8	9	10
Mean				579	574	594	600	541	585	578
Z-Score				3.96	3.72	4.69	4.99	2.04	4.23	3.92

DEUCE Program ($\mu_1 = 520$)

Trial	1	2	3	4	5	6	7	8	9	10
Mean				509	511	513	502	492	513	533
Z-Score				0.44	0.54	0.65	0.11	-0.41	0.65	1.63

3a. Using alpha of .01 instead of .05, how many times could you reject the null hypothesis for your results in **Exercise 1**? (How many times is Z > 2.326?)

	α = .05 (from #1)	α = .01
Reject for ACE Program (μ_1 = 580)		
Reject for DEUCE Program (μ_1 = 520)		

3b. What is the power for each of these tests? You can use the applet below to calculate power for the tests using alpha α = .01. (Set n = 25 and μ_0 = 500 for all tests ; use μ_1 = 580 for ACE and μ_1 = 520 for DEUCE). Remember to press 'Enter' after each change to the applet.

	α = .05 (from #1)	α = .01
Power for ACE Program (μ_1 = 580)	.991	
Power for DEUCE Program (μ_1 = 520)		

Cumulative Test: What affects Statistical Power?

Select True or False for each of the following questions.

If nothing else is changed, power is greater when…

	T	F	
C1.	☐	☐	The difference between the null and alternative population means is greater.
C2.	☐	☐	The standard deviation of the populations is greater.
C3.	☐	☐	The alpha error rate is changed from .01 to .05.
C4.	☐	☐	The sample size is changed from 30 to 40.

More challenging questions:

	T	F	
C5.	☐	☐	Power is always greater when the effect size is greater.
C6.	☐	☐	Power is always greater when the Type II error (i.e., beta error) is smaller.
C7.	☐	☐	To compute power, all I need to know is effect size, sample size, and alpha.

C8. Which of the following situations would yield the greatest power (assuming alpha and sample size are held constant)?

☐ Null mean = 500, Alternative mean = 510, Standard Deviation = 40

☐ Null mean = 500, Alternative mean = 540, Standard Deviation = 160

☐ Null mean = 500, Alternative mean = 520, Standard Deviation = 60

C9. Consider the shape of the sampling distributions for samples of size n = 4, n = 25, and n = 100. What happens to the sampling distribution of the sample mean when n is increased (assuming nothing else changes)?

◉ Sampling distribution becomes more disperse.

☐ Sampling distribution becomes less disperse.

☐ Sampling distribution remains the same.

C10. So far you have examined the effect of magnitude of difference between the null mean and the alternative mean, standard deviation, sample size, and alpha level on power. Which of the answers below best summarizes the effect of each on power?

- More power = large magnitude of difference, larger standard deviation, larger sample, larger alpha.

- More power = large magnitude of difference, smaller standard deviation, larger sample, smaller alpha.

- More power = large magnitude of difference, smaller standard deviation, larger sample, larger alpha.

- More power = smaller magnitude of difference, smaller standard deviation, larger sample, smaller alpha.