Type I & Type II error

DECISION	TRUE STATE: Ho	TRUE STATE: H1
Do not reject Ho	correct decision $p = 1-\alpha$	Type II error $p = \beta$
Reject Ho	Type I error $p = \alpha$	correct decision $p = 1-\beta$

- Type I error, α (alpha), is defined as the probability of rejecting a true null hypothesis
- Type II error, β (beta), is defined as the probability of failing to reject a false null hypothesis

Power

- Normally, no adverse consequences occur when we correctly fail to reject a null hypothesis
  - Declaring not guilty an innocent man → he is free to go
- Type I and II errors are mistakes we do not want to make
  - Letting a criminal go free (Type II)
  - Or worse, sending to jail an innocent man (Type I)
    - That's why we set alpha to 0.05
- On the other hand, the ability to convict a guilty person is essential to our justice system
  - Reject Ho, when Ho is false
- this ability, in statistics, is referred to as power

DECISION	TRUE STATE: not guilty	TRUE STATE: guilty
not guilty	correct decision $p = 1-\alpha$	Type II error $p = \beta$
guilty	Type I error $p = \alpha$	correct decision $p = 1-\beta$

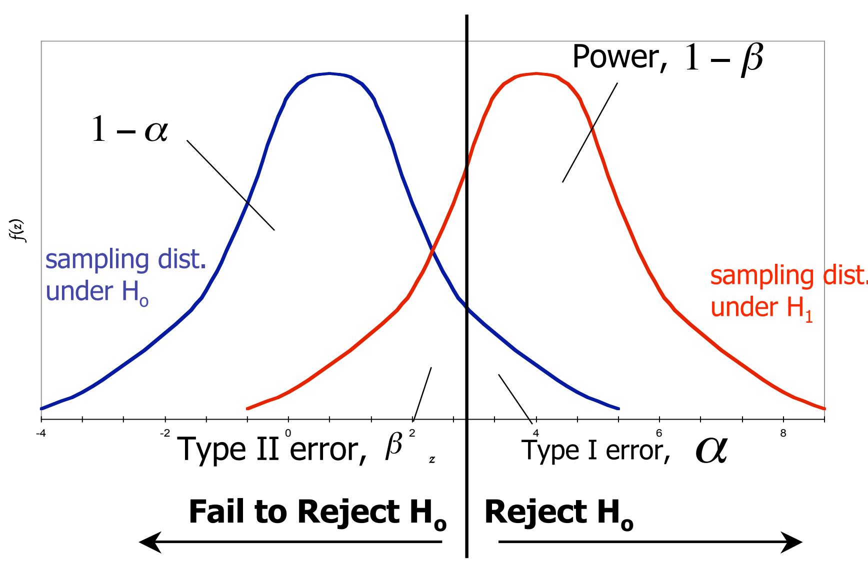
Power

Definition

- Power is the probability of correctly rejecting a false null hypothesis

DECISION	***TRUE STATE*** **Ho**	**H1**
Do not reject Ho	correct decision *p* =1-α	Type II error *p* = β
Reject Ho	Type I error *p* = α	correct decision *p* =1-β

Power & Errors



Power - Alpha & Beta

- β ➔ Type II Error: Fail to reject Ho even when H_1 is true.
  - Power = 1 - β
  - If we increase power we reduce β, we reduce the probability of getting a Type II Error
- α ➔ Criterion for the test, it tells you were to start rejecting Ho.
  - We usually set $\alpha = 0.05$
  - If we want a more stringent criterion,
    - α decreases
    - More difficult to reject Ho ➔ power decreases
    - Power decreases ➔ β increases (easier to make a Type II Error)

Since Power is the probability of correctly rejecting a false null hypothesis, It is to our best interest to increase power.

Ways of increasing Power

- make alpha larger
- use one-tailed rather than two tailed test
- decrease variance
  - increase sample size
  - better measures
- increase effect size

Magnitude of Power

- Strong effect
    - 0.00-0.15 or 0.85-1.00
- Moderate effect
    - 0.15-0.35 or 0.65-0.85
- Weak effect
    - 0.35-0.49 or 0.51-0.65

Calculating Power

1. Choose 1-tail or 2-tail test
2. Set alpha value
3. Select statistical test
4. Determine critical values to reject H_0
5. Calculate the probability of obtaining those values under a specified H_1
   – That is the power of the test under that version of H1

Calculating Power: Example 1

A researcher wants to identify the effects of sleep deprivation on test performance in his introductory statistics course.

A group of 19 students will perform under both of the following conditions:

- Sleep Deprivation: stay awake for 2 nights and days.
- Control: sleep normally for 2 nights and days.

After each 2-day period, the students are tested. The tests are scored.

a) Calculate the power of the experiment to test a moderate effect (0.80 or 0.20).

b) What's the probability of a Type II Error?

Calculating Power: Example 1 (cont)

1. **Choose 1-tail or 2-tail test**

   don't know which will have higher test scores ➔ 2-tail test

2. **Set alpha value**

   α= 0.05

3. **Select statistical test**

   Use sign test:
   - only two possible outcomes for hypothesis: + (E) or - (~E).
   - 19 students = 19 trials
     - exams are assumed to be independent (!!!!)
   - If there are only two outcomes, guessing the correct outcome is given by chance
     - p(E)=.50
     - p(~E)=.50
     - p+q=1
     - p and q are constant over trials

Calculating Power: Example 1 (cont)

4. Determine critical values to reject H_0

N # of +s	for P(E)=.50 p
0	0.0000
1	0.0000
2	0.0003
3	0.0018
4	0.0074
5	0.0222
6	0.0518
7	0.0961
8	0.1442
9	0.1762
10	0.1762
11	0.1442
12	0.0961
13	0.0518
14	0.0222
15	0.0074
16	0.0018
17	0.0003
18	0.0000
19	0.0000

Calculating Power: Example 1 (cont)

5. Calculate the probability of obtaining those values under a specified H_1

For a moderate effect 0.20 or 0.80 what is the probability of rejecting Ho?

We need to see probability for 4 or 15 pluses with P(E)=0.20 under a moderate effect.

N # of +s	for P(E)=.20 p
0	0.0115
1	0.0576
2	0.1369
3	0.2054
4	0.2182
5	0.1746
6	0.1091
7	0.0545
8	0.0222
9	0.0074
10	0.0020
11	0.0005
12	0.0001
13	0.0000
14	0.0000
15	0.0000
16	0.0000
17	0.0000
18	0.0000
19	0.0000

What's the probability of a Type II Error?
β= 1 - power

Calculating Power: Example 2 (from Pagano Prob. 11.3)

A TV program is believed to cause violence among teenagers. You want to test this belief in a scientific manner, you start by obtaining a random sample of 15 teens from the local high school.

Each teen is run in both the experimental and control condition.

Experimental Condition: Watch TV program for a period of 3 months and record number of violent acts within those 3 months.

Control Condition: Do not watch TV program for a period of 3 months and record number of violent acts within those 3 months.

a) Calculate the power of the experiment to test a moderate effect of 0.70 in the direction of the hypothesis.

b) What's the probability of a Type II Error?

Calculating Power: Example 2 (cont)

1. **Choose 1-tail or 2-tail test**

   increase in violence ➔ 1-tail test

2. **Set alpha value**

   $\alpha = 0.05$

3. **Select statistical test**

   Use sign test:
   - only two possible outcomes for hypothesis: + (E) or - (~E).
   - 15 subjects = 15 trials
     - Actions among subjects are assumed to be independent (!!!!)
   - If there are only two outcomes, guessing the correct outcome is given by chance
     - p(E)=.50
     - p(~E)=.50
     - p+q=1
     - p and q are constant over trials

Calculating Power: Example 1 (cont)

4. Determine critical values to reject H_0

N # of +s	for P(E)=.50 p
0	0.0000
1	0.0005
2	0.0032
3	0.0139
4	0.0417
5	0.0916
6	0.1527
7	0.1964
8	0.1964
9	0.1527
10	0.0916
11	0.0417
12	0.0139
13	0.0032
14	0.0005
15	0.0000

Calculating Power: Example 2 (cont)

5. Calculate the probability of obtaining those values under a specified H_1

For a moderate effect of 0.70 what is the probability of rejecting Ho?

We need to see probability for 12 pluses with P(E)=0.70 under a moderate effect.

NOTE: Tables - P(~E)=0.30

What's the probability of a Type II Error?

N # of +s	for P(~E)=.30 p
0	0.0047
1	0.0305
2	0.0916
3	0.1700
4	0.2186
5	0.2061
6	0.1472
7	0.0811
8	0.0348
9	0.0116
10	0.0030
11	0.0006
12	0.0001
13	0.0000
14	0.0000
15	0.0000

Sampling Distributions

- We draw inferences about population parameters from sample statistics
  - Sample proportion approximates population proportion
  - Sample mean approximates population mean
  - Sample variance (using n-1) approximates population variance
  - Etc.
- Statistics vary from one sample to the next
  - If a statistic is unbiased, then on average over many samples, it will equal the population parameter
  - There will be variability around that average
  - **The distribution of a statistic is the sampling distribution**

Sampling Distributions

Definition

The Sampling Distribution gives all the values a statistic can take, along with the probability of getting each value if sampling is random from the null hypothesis population.

Example:

The sampling distribution of the mean scores of a PSYC-200* exam can tell us all the possible mean scores we can get from random samples and how probable those mean scores can be.

* Population of students who take PSYC-200

Find the sampling distribution of the mean from the following population where the sample size is 2.

- Sampling should be done with replacement.

Sample	Elements		Mean
1	20	20	20
2	20	30	25
3	20	40	30
4	20	50	35
5	30	20	25
6	30	30	30
7	30	40	35
8	30	50	40
9	40	20	30
10	40	30	35
11	40	40	40
12	40	50	45
13	50	20	35
14	50	30	40
15	50	40	45
16	50	50	50

We need to get all possible combinations we could get by sampling from the population.

Next, we need to find the probability for each sample mean.

Sample	Elements		Mean
1	20	20	20
2	20	30	25
3	20	40	30
4	20	50	35
5	30	20	25
6	30	30	30
7	30	40	35
8	30	50	40
9	40	20	30
10	40	30	35
11	40	40	40
12	40	50	45
13	50	20	35
14	50	30	40
15	50	40	45
16	50	50	50

$$p(20) = \frac{1}{16}$$

$$p(25) = \frac{2}{16}$$

$$p(30) = \frac{3}{16}$$

$$p(35) = \frac{4}{16}$$

$$p(40) = \frac{3}{16}$$

$$p(45) = \frac{2}{16}$$

$$p(50) = \frac{1}{16}$$

What's the probability that the next sample we get has a mean of 35?

We could draw the sampling distribution based on the probability scores.