Lecture 28/Chapters 22 & 23

Hypothesis Tests

- Variable Types and Appropriate Tests
- Choosing the Right Test: Examples
- Example: Reviewing Chi-Square
- Type I and Type II Error

Choosing the Right Test *(Review)*

Type of test depends on variable types:

- 1 categorical: z test about population proportion
- 1 measurement (quan) [pop sd known or sample large]: z test about mean
- 1 measurement (quan) [pop sd unknown & sample small]: t test about mean
- 1 categorical (2 groups)+ 1 quan: two-sample z or t
- 2 categorical variables: chi-square test (done in Chapter 13)

Null and Alternative Hypotheses *(Review)*

For a test about a single mean,

- **Null hypothesis:** claim that the population mean equals a proposed value.
- **Alternative hypothesis:** claim that the population mean is greater, less, or not equal to a proposed value.

  An alternative formulated with ≠ is **two-sided**; with > or < is **one-sided**.

Testing Hypotheses About a Population

1. Formulate hypotheses
   - about single proportion or mean or two means (alternative can have < or > or ≠ sign)
   - about relationship using chi-square: null hyp states two cat. variables are not related; alt states they are.
2. Summarize/standardize data.
3. Determine the P-value. (2-sided is twice 1-sided)
4. Make a decision about the population: believe alt if P-value is small; otherwise believe null.

For practice, we'll consider a variety of examples. In each case we'll formulate appropriate hypotheses and state what type of test should be run.

Example: *Smoking and Education (#1 p. 427)*

- **Background**: Consider years of education for mothers who smoke compared with those who don't, in sample of 400 mothers, to decide if one group tends to be more educated.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: z test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: z test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: t test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample z or t
  5. 2 categorical variables: chi-square test
- **Response:** _____

Example: *Test about Smoking and Education*

- **Background**: Consider years of education for mothers who smoke compared with those who don't, in sample of 400 mothers, to decide if one group tends to be more educated.
- **Question:** What hypotheses and test are appropriate?
- **Response:**

Null: ____

Alt: ____

Do _______________ [large samples] test to compare _____________

Alternative is___________ because no initial suspicion was expressed about a specific group being better educated.

Example: *ESP? (Case Study 22.1 p. 425)*

- **Background**: A subject in an ESP experiment chooses each time from 4 targets the one which he/she believes is being "sent" by extrasensory means. Researchers want to determine if the subject performs significantly better than one would by random guessing.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: z test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: z test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: t test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample z or t
  5. 2 categorical variables: chi-square test
- **Response:** ____

Example: *Test about ESP*

- **Background**: A subject in an ESP experiment chooses each time from 4 targets the one which he/she believes is being "sent" by extrasensory means. Researchers want to determine if the subject performs significantly better than one would by random guessing.
- **Question:** What hypotheses and test are appropriate?
- **Response:**

Null: population proportion correct ________

Alt: population proportion correct _________

Do___ test about _______________________

Example: *Calcium for PMS (#3-4 p. 428)*

- **Background**: We want to compare change in severity of PMS symptoms (before minus after, measured quantitatively) for 231 women taking calcium vs. 235 on placebo to see if calcium helps.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: *z* test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: *z* test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: *t* test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample *z* or *t*
  5. 2 categorical variables: chi-square test
- **Response:** ____

Example: *Test about Calcium for PMS*

- **Background**: We want to compare change in severity of PMS symptoms (before minus after, measured quantitatively) for 231 women taking calcium vs. 235 on placebo to see if calcium helps.
- **Question:** What hypotheses and test are appropriate?
- **Response:**

Null: mean symptom change (calc)__mean symptom change (placebo)

Alt: mean symptom change (calc)__mean symptom change (placebo)

Do _____________ [large samples] test to compare means

Alternative is__________ because we hope or suspect that the calcium group will show more symptom improvement.

As always, our hypotheses refer to the___________, not the________

Example: *Incubators, Claustrophobia (6b p.428)*

- **Background**: We want to see if placing babies in an incubator during infancy can lead to claustrophobia in adult life.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: *z* test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: *z* test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: *t* test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample *z* or *t*
  5. 2 categorical variables: chi-square test
- **Response:** ____

Example: *Test about Incubators, Claustrophobia*

- **Background**: We want to see if placing babies in an incubator during infancy can lead to claustrophobia in adult life.
- **Question:** What hypotheses and test are appropriate?
- **Response:**

Null: there is___relationship between incubation and claustrophobia

Alt: there is___relationship between incubation and claustrophobia

Do ____________test.

Alternative is general (2-sided) because __________doesn't let us specify our initial suspicions in a particular direction.

Example: *Training Program, Scores (#7 p.446)*

- **Background**: We want to see if a training program helps raise students' scores. For each student, researchers record the increase (or decrease) in the scores, from pre-test to post-test.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: z test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: z test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: t test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample z or t
  5. 2 categorical variables: chi-square test
- **Response:** ______________________________

Note: 2-sample design would be better, to avoid placebo effect.

Example: *Test about Training Program, Scores*

- **Background**: We want to see if a training program helps raise students scores. For each student, researchers record the increase (or decrease) in the scores, from pre-test to post-test.
- **Question:** What hypotheses and test are appropriate?
- **Response:**

Null: population mean increase___

Alt: population mean increase___

Call it a ______________ (not sure if sample is large enough to use z) based on a matched-pairs design (see page 88).

Alternative is__________ because the training program is supposed to help.

Note: As always, our hypotheses refer to population values. It's not enough to simply exhibit an increase in sample scores; the increase must be **statistically significant**.

Example: *Terrorists' Religion: Discrimination?*

- **Background**: We want to see if Catholics were discriminated against, based on a table of religion and acquittals for persons charged with terrorist offenses in Northern Ireland in 1991.
- **Question:** Which of the 5 situations applies?
  1. 1 categorical: z test about population proportion
  2. 1 measurement (quan) [pop sd known or sample large]: z test about mean
  3. 1 measurement (quan) [pop sd unknown & sample small]: t test about mean
  4. 1 categorical (2 groups) + 1 quan: two-sample z or t
  5. 2 categorical variables: chi-square test
- **Response:** ____

Chi-Square Test *(Review)*

We learned to use chi-square to test for a relationship between two categorical variables.

1. Null hypothesis: the two variables are not related
   alternative hypothesis: the two variables are related
2. Test stat = chi-sq = sum of $\frac{(\text{observed count}-\text{expected count})^2}{\text{expected count}}$
3. P-value= probability of chi-square this large, assuming the two variables are not related. For a 2-by-2 table, chi-square > 3.84 ↔ P-value < 0.05.
4. If the P-value is small, conclude the variables are related. Otherwise, we have no convincing evidence of a relationship.

Note: Next lecture we'll do another example of a chi-square test.

Example: *Chi-Square Review: Discrimination?*

- **Background**: Table for religion and trial outcome:

Observed	Acquitted	Convicted	Total
Protestant	8	7	15
Catholic	27	38	65
Total	35	45	80

- **Question:** What do we conclude?
- **Response:** First formulate hypotheses.

Null: there is___relationship between religion and trial outcome

Alt: there is___relationship between religion and trial outcome

Are Variables in a 2×2 Table Related?

1. Compute each expected count $= \dfrac{\textit{Column total} \times \textit{Row total}}{\textit{Table total}}$
2. Calculate each component $= \dfrac{(\text{observed - expected})^2}{\text{expected}}$
3. Find chi-square = sum of $\dfrac{(\text{observed - expected})^2}{\text{expected}}$
4. If chi-square > 3.84, there is a statistically significant relationship. Otherwise, we don't have evidence of a relationship.

Example: *Religion & Acquittal Related?*

- **Background**: Two-way table for religion and trial outcome:

Observed	Acquitted	Convicted	Total
Protestant	8	7	15
Catholic	27	38	65
Total	35	45	80

- **Question:** What counts would we expect if there were no relationship?
- **Response:** Expect…
  - ________________________ Protestants to be acquitted
  - ________________________ Catholics to be acquitted
  - ________________________ Protestants to be convicted
  - ________________________ Catholics to be convicted

Example: *Religion & Acquittal (continued)*

- **Background**: Observed and Expected Tables:

Obs	Acquitted	Convicted	Total
Prot	8	7	15
Cath	27	38	65
Total	35	45	80

Exp	Acquitted	Convicted	Total
Prot	6.56	8.44	15
Cath	28.44	36.56	65
Total	35	45	80

- **Question:** Find components & chi-square; conclude?
- **Response:** chi-square =

=0.32 + 0.25 + 0.07 + 0.06 = 0.70

The relationship is ______________________________ We _____ have convincing evidence of a relationship (discrimination).

Example: *HIV Test (Review)*

- **Background**: In a certain population, the probability of HIV is 0.001. The probability of testing positive is 0.98 if you have HIV, 0.05 if you don't.
- **Questions:** What is the probability of having HIV and testing positive? Overall prob of testing positive? Probability of having HIV, given you test positive?
- **Response:** To complete the tree diagram, note that probability of not having HIV is 0.999. The probability of testing negative is 0.02 if you have HIV, 0.95 if you don't.

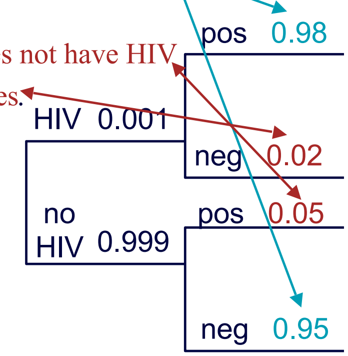
Example: *HIV Test (Review)*

Possible correct conclusions:

- positive test when someone has HIV
- negative test when someone does not have HIV

Possible incorrect conclusions:

- positive test when someone does not have HIV
- negative test when someone does.



Two Types of Error

Decision→ Actuality↓	Healthy (don't reject null hyp)	Diseased (reject null hyp)
Healthy (null hyp true)	Correct (prob= specificity=0.95)	Incorrect: false positive= Type I Error (prob=0.05)
Diseased (alt hyp true)	Incorrect: false negative= Type II Error (prob=0.02)	Correct (prob=sensitivity=0.98)

If we decide in advance to use 0.05 as our cut-off for a small *P*-value, then 0.05 will be our probability of a Type I Error. The probability of a Type II Error can be specified only if we happen to know what is true in actuality (observed in the long run?).