Statistical Inference

Kosuke Imai

Department of Politics
Princeton University

Fall 2011

What is Statistical Inference?

- READING: FPP Chapter 19
- Guessing what you do not observe from what you do observe
- Start with the probability model with some unknown parameters θ
- Use the data to estimate the parameters $\hat{\theta}$
- Compute uncertainty of your estimates

- Sampling distribution of $\hat{\theta}$
- Estimation error: $\hat{\theta} - \theta_0$ where θ_0 is the true value
- Bias: $B = \mathbb{E}(\hat{\theta}) - \theta_0$
- Efficiency: $\mathbb{V}(\hat{\theta})$
- Mean Squared Error (MSE): $\mathbb{E}(\hat{\theta} - \theta_0)^2 = B^2 + \mathbb{V}(\hat{\theta})$
- Consistency: $\hat{\theta} \to \theta_0$ as n increases

The 1936 Literary Digest Poll

- Source: automobile registration, telephone books, etc.
- Initial sample size: over 10 million straw vote ballots
- Final sample size: over 2.3 million returned
- The young Gallup used 50,000 respondents

	FDR's vote share
Literary Digest	43
George Gallup	56
Gallup's prediction of LD	44
Actual Outcome	62

The 1948 Election

The Polling Disaster

	Truman	Dewey	Thurmond	Wallace
Crossley	45	50	2	3
Gallup	44	50	2	4
Roper	38	53	5	4
Actual Outcome	50	45	3	2

- Quota sampling
- fixed quota of certain respondents for each interviewer
- sample resembles the population w.r.t. these characteristics
- potential selection bias

Survey Sampling and Potential Bias

- Probability sampling to ensure representativeness
- Definition: every unit in the population has a known non-zero probability of being selected

Population	Sample	Potential Bias
Inferential population		
$\downarrow$		Frame bias
Frame population $\rightarrow$	Sample	Sampling bias
	$\downarrow$	Unit nonresponse
	Respondents	
	$\downarrow$	Item nonresponse
	Completed items	
		Response bias

Modern Election Polls

- Some 2004 national polls:

	registered voters		likely voters	
Polls	Bush	Kerry	Bush	Kerry
CNN/USA Today (Oct. 14-16)	49	46	52	44
Newsweek (Oct. 14-15)	48	46	50	44
New York Times (Oct. 14-17)	45	45	47	45
Time (Oct. 14-15)	46	46	48	47
ABC News (Oct. 14-17)	49	46	50	47

- Sample size $\approx 1,200$
- margin of error $\approx$ 3 percentage points
- What is margin of error?
- Why can't pollsters agree?

Estimating the Population Proportion

- READINGS: FPP Chapters 20, 21, and 23; A&F 4.4–4.5, 5.1–5.4

- Assume a random sample from a very large (infinite) population
- Assume everyone is equally likely to be sampled
- Survey with n respondents: $X_i = 1$ if respondent i supports Obama, $X_i = 0$ otherwise

- Population proportion (parameter): $\mathbb{E}(X) = \Pr(X_i = 1) = p$
- Sample proportion (statistic): $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

- Unbiasedness: $\mathbb{E}(\overline{X}) = p$
- Consistency (Law of Large Numbers): $\overline{X} \to p$ as n increases

Sampling Distribution and Standard Error

- Sampling variance:

$$\mathbb{V}(\overline{X}) = \frac{p(1-p)}{n}$$

- Standard error $=$ *Estimated* std. dev. of the sampling distribution:

$$\text{s.e.} = \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}$$

- For a large sample size,
  1. By law of large numbers,

$$\overline{X} \approx p \quad \text{and} \quad \text{s.e.} \approx \sqrt{p(1-p)/n}$$

  2. By the central limit theorem,

$$\overline{X} \sim \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$$

- Use these results to make approximate inference

Confidence Intervals

- $(1-\alpha) \times 100\%$ (asymptotic) confidence intervals:

$$\mathrm{CI}_\alpha \;=\; [\overline{X} - z_{\alpha/2}\,\text{s.e.},\; \overline{X} + z_{\alpha/2}\,\text{s.e.}]$$

  where $z_{\alpha/2}$ is called the critical value
- $\Pr(Z > z_{\alpha/2}) = \alpha/2$ and $Z \sim \mathcal{N}(0, 1)$
  1. $\alpha = 0.01$ gives $z_{\alpha/2} = 2.58$
  2. $\alpha = 0.05$ gives $z_{\alpha/2} = 1.96$
  3. $\alpha = 0.10$ gives $z_{\alpha/2} = 1.64$
- Other names: confidence bands, error bands, etc.
- Be careful about the interpretation!!
  1. Probability that the true value is in a *particular* confidence interval is either 0 or 1
  2. Confidence intervals are *random*, while the truth is *fixed*

An Example: Obama's Approval Rate

- CNN Poll: Do you approve or disapprove the way Barack Obama is handling his job as president?
- October 30 – November 1, 2009: Approve 54%, Disapprove 45%, No opinion 1%
- 1,018 adult Americans
- Parameter: Population proportion of Americans who approve Obama
- Statistic: Sample proportion of those who approve Obama
- Estimate 54%, Standard error 0.02, 95% CI $[0.500, 0.583]$

Margin of Error in Polls

- In a close election

$$\text{MoE} = \pm z_{0.025}\ \text{s.e.} \approx \pm 1.96 \times \sqrt{\frac{0.5 \times 0.5}{n}} \approx \pm\frac{1}{\sqrt{n}}$$

- A typical poll: $n = 1,200$ and $\text{MoE} = \pm 2.9$ percentage points
- Sample size calculation:

$$n \approx \frac{1}{\text{MoE}^2}$$

- More generally,

$$n = \frac{1.96^2 p(1-p)}{\text{MoE}^2}$$

- A larger sample is needed when p is closer to 0.5
- What is the minimal sample size required when $p = 0.3$ and $\text{MoE} = 0.03$? About 900

Another Example: Margin of Victory

- Recall the Obama-McCain example
- Margin of victory: $\delta = 2p - 1$
- Estimate: $\hat{\delta} = 2\overline{X} - 1$
- Standard error: $\text{s.e.} = 2\sqrt{\overline{X}(1 - \overline{X})/n}$
- Approximate sampling distribution of the Z-statistic

$$Z = \frac{\hat{\delta} - \delta}{s.e.} \overset{\text{approx.}}{\sim} \mathcal{N}(0, 1)$$

- $(1 - \alpha)$% confidence intervals: $[\hat{\delta} - \text{s.e.} \times z_{\alpha/2}, \ \hat{\delta} + \text{s.e.} \times z_{\alpha/2}]$
- 3 percentage point lead with the sample size of 1,200: $\text{s.e.} = 0.029$ and 95% CI $= [-0.03, 0.09]$
- Uncertainty is larger than the estimated support for one candidate

Inference with Normally Distributed Data

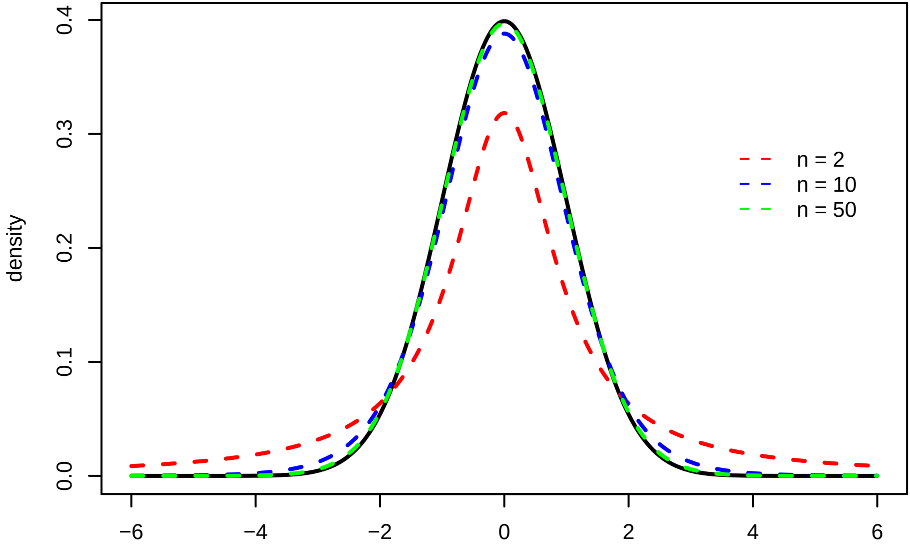
- If $X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$, then $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ *without* the CLT
- Standard error: $\text{s.e.} = S_X/\sqrt{n}$
- Moreover, *without* the CLT

$$t\text{–statistic} = \frac{\overline{X} - \mu}{\text{s.e.}} \overset{\text{exactly}}{\sim} t_{n-1}$$

  where t_{n-1} is the t distribution with $n - 1$ degrees of freedom
- Use t_{n-1} (rather than $\mathcal{N}(0, 1)$) to obtain the critical value for exact confidence intervals

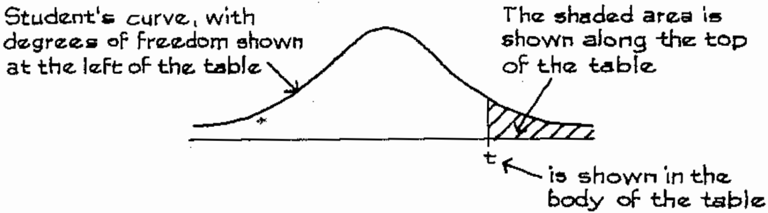
- As n increases, t_{n-1} approaches to $\mathcal{N}(0, 1)$
- CI based on t distribution is wider than CI based on normal
- More conservative inference

Student's *t* Distribution



t Table (FPP A–105)

A *t*-TABLE



Degrees of freedom	*25%*	*10%*	*5%*	*2.5%*	*1%*	*0.5%*
1	1.00	3.08	6.31	12.71	31.82	63.66
2	0.82	1.89	2.92	4.30	6.96	9.92
3	0.76	1.64	2.35	3.18	4.54	5.84
4	0.74	1.53	2.13	2.78	3.75	4.60
5	0.73	1.48	2.02	2.57	3.36	4.03

Overview of Statistical Hypothesis Testing

- READINGS: FPP Chapters 26 and 29, A&F 6.1–6.6
- Probabilistic "Proof by contradiction": Assume the negation of the proposition, and show that it leads to the contradiction

1. Construct a null hypothesis (H_0) and its alternative (H_1)
2. Pick a test statistic T
3. Figure out the sampling distribution of T under H_0 (reference distribution)
4. Is the observed value of T likely to occur under H_0?
   - Yes – Retain H_0
   - No – Reject H_0

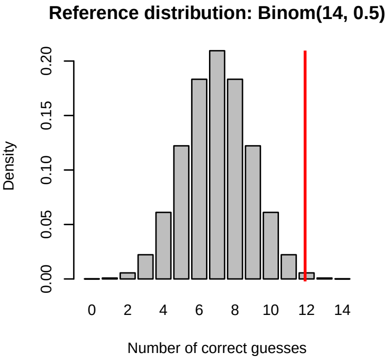
Paul the Octopus

- 2010 World Cup
  - Group: Germany vs Australia
  - Group: Germany vs Serbia
  - Group: Ghana vs Germany
  - Round of 16: Germany vs England
  - Quarter-final: Argentina vs Germany
  - Semi-final: Germany vs Spain
  - 3rd place: Uruguay vs Germany
  - Final: Netherlands vs Spain

- Question: Did Paul the Octopus get lucky?
- Null hypothesis: Paul is randomly choosing winner
- Test statistics: Number of correct answers
- Reference distribution: $\text{Binomial}(8, 0.5)$
- The probability that Paul gets them all correct: $\frac{1}{2^8} \approx 0.004$
- Tie is possible in group rounds: $\frac{1}{3^3} \times \frac{1}{2^5} \approx 0.001$

More Data about Paul

- UEFA Euro 2008
  - Group: Germany vs Poland
  - ~~Group~~: Croatia vs Germany
  - Group: Austria vs Germany
  - Quarter-final: Portugal vs Germany
  - Semi-final: Germany vs Turkey
  - ~~Final~~: Germany vs Spain
- A total of 14 matches
- 12 correct guesses



- *p*-value: Probability that under the null you observe something at least as extreme as what you actually observed
- $\Pr(\{12, 13, 14\}) \approx 0.001$
- R: `pbinom(12, size = 14, prob = 0.5, lower.tail = FALSE)`

Paul's Rival, Mani the Parakeet

- 2010 World Cup
  - Quarter-final: Netherlands vs Brazil
  - Quarter-final: Uruguay vs Ghana
  - Quarter-final: Argentina vs Germany
  - Quarter-final: Paraguay vs Spain
  - ~~Semi-final~~: Uruguay vs Netherlands
  - Semi-final: Germany vs Spain
  - ~~Final~~: Netherlands vs Spain

- Mani did pretty good: p-value is 0.0625

- Danger of multiple testing
- Take 10 animals with no forecasting ability. What is the chance of getting p-value less than 0.05 at least once?

$$1 - 0.95^{10} \approx 0.4$$

- If you do this with enough animals, you will find another Paul

Hypothesis Testing for Proportions

1. Hypotheses – $H_0 : p = p_0$ and $H_1 : p \neq p_0$
2. Test statistic: $\overline{X}$
3. Under the null, by the central limit theorem

$$Z\text{-statistic} \;=\; \frac{\overline{X} - p_0}{\text{s.e.}} \;=\; \frac{\overline{X} - p_0}{\sqrt{p_0(1 - p_0)/n}} \overset{\text{approx.}}{\sim} \mathcal{N}(0, 1)$$

4. Is Z_{obs} unusual under the null?
    - Reject the null when $|Z_{obs}| > z_{\alpha/2}$
    - Retain the null when $|Z_{obs}| \leq z_{\alpha/2}$

- The level (size) of the test: $\Pr(\text{rejection} \mid H_0) \;=\; \alpha$
- Duality with confidence intervals:
    - Reject the null $\Longleftrightarrow$ p_0 not in CI_α
    - Retain the null $\Longleftrightarrow$ p_0 in CI_α
- When X_i is normally distributed, use t-statistic and obtain the critical value using Student's t distribution

p-value

- (two-sided) p-value$=\Pr(Z > |Z_{obs}|) + \Pr(Z < -|Z_{obs}|)$
- One sided alternative hypothesis: $H_1 : p > p_0$ or $p < p_0$
- one-sided p-value$=\Pr(Z > Z_{obs})$ or $\Pr(Z < Z_{obs})$
- Use `pnorm()` or `pt()`

- p-value is the probability, computed under H_0, of observing a value of the test statistic at least as extreme as its observed value
- A smaller p-value presents stronger evidence against H_0
- p-value less than α indicates statistical significance $\leftrightarrow$ α-level test

- p-value is NOT the probability that H_0 (H_1) is true (false)
- The statistical significance indicated by the p-value does not necessarily imply scientific significance

Back to the Polling Examples

1. Obama's approval rate
   - $H_0 : p = 0.5$ and $H_1 : p \neq 0.5$
   - $\alpha = 0.05$ level test
   - $\overline{X} = 0.54$ and $n = 1018$
   - $Z_{obs} = (0.54 - 0.5)/\sqrt{0.5 \times 0.5/1018} = 2.55 > z_{0.025} = 1.96$
   - p-value $= 0.005 \times 2 = 0.010$
   - Reject the null
2. Obama's margin of victory
   - $H_0 : \delta = 0$ and $H_1 : \delta > 0$
   - $\alpha = 0.1$ level test
   - $\hat{\delta} = 0.03$ and $n = 1200$
   - $Z_{obs} = 0.03/\sqrt{1/1200} = 1.04 < z_{0.1} = 1.28$
   - p-value $= 0.149$
   - Retain the null

Error and Power of Hypothesis Test

- Two types of errors:

	Reject H_0	Retain H_0
H_0 is true	Type I error	Correct
H_0 is false	Correct	Type II error

- Hypothesis tests control the probability of Type I error, which is equal to the level of tests or α
- They do not control the probability of Type II error
- Tradeoff between the two types of error
- A large p-value can occur either because H_0 is true or because H_0 is false but the test is not powerful

- Level of test: probability that the null is rejected when it is true
- Power of test: probability that a test rejects the null
- Typically, we want a most powerful test given the level

Power Analysis

- Null hypotheses are often uninteresting
- But, hypothesis testing may indicate the strength of evidence for or against your theory
- Power analysis: What sample size do I need in order to detect a certain departure from the null?
- Power $= 1 - \Pr(\text{Type II error})$
- Three steps
  1. Suppose $\mu = \mu^*$ which implies $\overline{X} \sim \mathcal{N}(\mu^*, \mathbb{V}(X)/n)$
  2. Calculate the rejection probability noting that we reject $H_0 : \mu = \mu_0$ if $|\overline{X}| > \mu_0 + z_{\alpha/2} \times \text{s.e.}$
  3. Find the smallest n such that this rejection probability equals a pre-specified level

One-Sided Case

- $H_0 : p = p_0$ and $H_a : p > p_0$
- $\overline{X} \sim \mathcal{N}(p^*, p^*(1-p^*)/n)$
- Reject H_0 if $\overline{X} > p_0 + z_{\alpha/2} \times \sqrt{p_0(1-p_0)/n}$

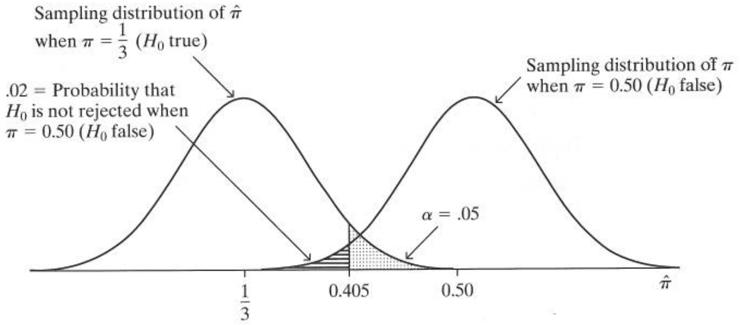


FIGURE 6.11: Calculation of P(Type II Error) for Testing H_0: $\pi = 1/3$ against H_a: $\pi > 1/3$ at $\alpha = 0.05$ Level, when True Proportion is $\pi = 0.50$. A Type II error occurs if $\hat{\pi} < 0.405$, since then P-value >0.05 even though H_0 is false.

Power Analysis for the Margin of Victory Example

- $H_0 : \delta = 0$ and $H_1 : \delta \neq 0$
- How large the sample size must be in order to reject H_0 at the 95% confidence level with probability 0.9 when $\delta = \pm 0.02$?
- Sampling distribution of $\hat{\delta}$ when $\delta = 0.02$:

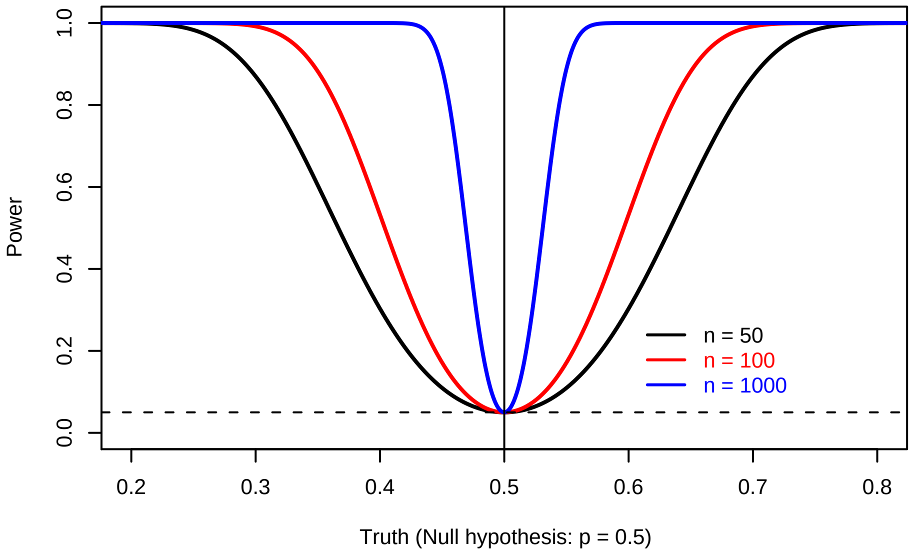
$$\hat{\delta} \sim \mathcal{N}(0.02,\ 4(0.51 \times 0.49)/n)$$

- Reject $H_0 : \delta = 0$ if $|\hat{\delta}| > 1.96/\sqrt{n}$
- Find the smallest value of n such that

$$\Pr(\text{Reject } H_0) = \Pr(\hat{\delta} > 1.96/\sqrt{n}) + \Pr(\hat{\delta} < -1.96/\sqrt{n}) = 0.9$$

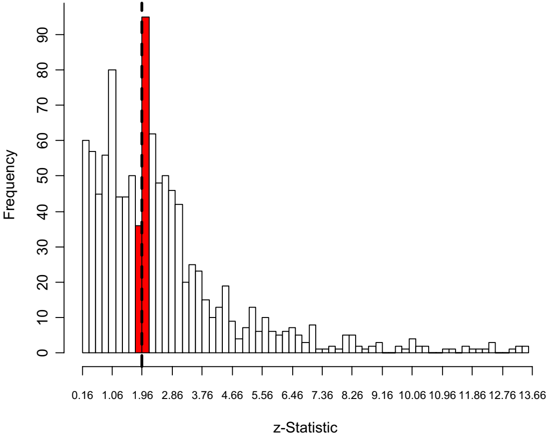
- How do you calculate $\Pr(\hat{\delta} > 1.96/\sqrt{n})$ using the standard normal table or R?

Power Function



Pitfalls of Hypothesis Testing

- Reporting CIs is typically better than conducting hypothesis tests
- Danger of multiple testing
- Publication bias (Gerber and Malhotra, *QJPS* 2008):



Comparison of Two Samples

- Comparison of two groups is more interesting
- Gender difference in public opinion about war
- Difference between treatment and control groups in experiments

- Causal inference with randomized experiments
- Back to the GOTV example
- The 2006 Michigan August primary experiment
- Treatment Group: postcards showing their own and their neighbors' voting records
- Control Group: received nothing

Social Pressure Experiment

- Turnout rate: $\overline{X}_T = 0.37$, $\overline{X}_C = 0.30$,
- Sample size: $n_T = 360$, $n_C = 1890$
- Estimated average treatment effect:

$$\text{est.} \;=\; \overline{X}_T - \overline{X}_C \;=\; 0.07$$

- Standard error:

$$\text{s.e.} \;=\; \sqrt{\frac{\overline{X}_T(1-\overline{X}_T)}{n_T} + \frac{\overline{X}_C(1-\overline{X}_C)}{n_C}} \;=\; 0.028$$

- 95% Confidence intervals:

$$[\text{est.} - \text{s.e.} \times z_{0.025},\; \text{est.} + \text{s.e.} \times z_{0.025}] \;=\; [0.016,\; 0.124]$$

Social Pressure Example (Continued)

- Two-sample test
  - $H_0 : p_T = p_C$ and $H_1 : p_T \neq p_C$.
  - Reference distribution: $\mathcal{N}\left(0, \frac{p(1-p)}{n_T} + \frac{p(1-p)}{n_C}\right)$
  - p-value: 0.010
- Power calculation:
  - $p_T = 0.37$ and $p_C = 0.30$
  - Two-sample test at the 5% significance level
  - Equal group size: $n_T = n_C$
  - If $n = 1000$, what is the power of the test?

Another Example: The List Experiment

- Survey methodology for studying sensitive questions
- Minimizing nonresponse and social desirability bias
- The 1991 National Race and Politics Survey
- Randomize the sample into the treatment and control groups
- The script for the control group

```
Now I'm going to read you three things that sometimes
make people angry or upset. After I read all three,
just tell me HOW MANY of them upset you. (I don't
want to know which ones, just how many.)

(1) the federal government increasing the tax on
gasoline;
(2) professional athletes getting million-dollar-plus
salaries;
(3) large corporations polluting the environment.
```

Another Example: The List Experiment

- Survey methodology for studying sensitive questions
- Minimizing nonresponse and social desirability bias
- The 1991 National Race and Politics Survey
- Randomize the sample into the treatment and control groups
- The script for the treatment group

```
Now I'm going to read you four things that sometimes
make people angry or upset. After I read all four,
just tell me HOW MANY of them upset you. (I don't
want to know which ones, just how many.)

(1) the federal government increasing the tax on
gasoline;
(2) professional athletes getting million-dollar-plus
salaries;
(3) large corporations polluting the environment;
(4) a black family moving next door to you.
```

Analyzing List Experiments

- Estimating the proportion of white respondents who say they would be upset by "a black family moving next door to you":
  - Sample sizes: $n_T = 624$ and $n_C = 589$
  - Mean responses: $\overline{X}_T = 2.20$ and $\overline{X}_C = 2.13$
  - Standard deviations of responses: $S_T = 0.92$ and $S_C = 0.80$
  - How do you drive the 95% confidence intervals?

- Conducting a statistical hypothesis test:
  - Null hypothesis: Nobody is upset by it
  - How would you conduct a statistical test whose level is 5%?
  - What is the *p*-value?

- Conducting a power analysis:
  - Assume that the standard deviation of responses is equal to the overall standard deviation, which is 0.87
  - What is the minimal sample size that allows us to detect 5% of white respondents are being upset by the statement?

Statistical Inference with Regression

- Recall the model:

$$Y_i = \alpha + \beta X_i + \epsilon_i,$$

  where $\mathbb{E}(\epsilon_i) = 0$ and $\mathbb{V}(\epsilon_i) = \sigma^2$
- Key Assumptions:
  1. Exogeneity: the mean of ϵ_i does not depend on X_i
  2. Homoskedasticity: the variance of ϵ_i does not depend on X_i
- Heteroskedasticity is easy to fix but endogeneity is not
- When do we know the exogeneity assumption hold?
- Under exogeneity, $\hat{\alpha}$ and $\hat{\beta}$ are unbiased
- Under the two assumptions, standard errors can be obtained
- Usual hypothesis test: $H_0 : \beta = 0$
- Next week's precept: `summary(lm(y ~ x, data = data))`

Recall "Looks and Politics"

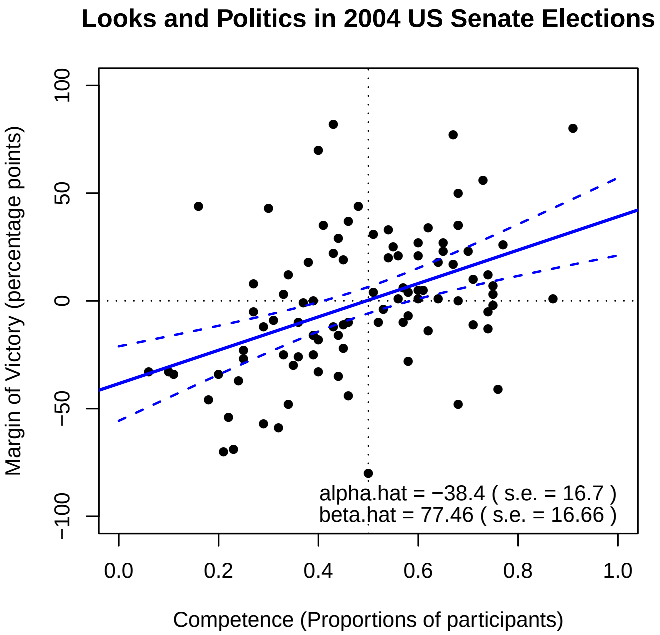
Todorov et al. Science

R Regression Summary Output

```
Call:
lm(formula = diff ~ comp, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-80.333 -16.756  -3.967  17.744  87.089

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -38.399      8.708  -4.410 2.78e-05 ***
comp          77.465     16.663   4.649 1.10e-05 ***
--
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 30.09 on 93 degrees of freedo
Multiple R-squared: 0.1886, Adjusted R-squared: 0.1798
```

Expected and Predicted Values

- Interpretation of β: the average increase in Y_i associated with one unit increase in X_i
- Expected value: the average outcome given X_i
- Predicted value: the prediction of the outcome given X_i
- Point estimate: $\hat{\alpha} + \hat{\beta} x$
- Standard error for expected value: $\sqrt{\mathbb{V}(\hat{\alpha} + \hat{\beta} x)}$
- Standard error for predicted value: $\sqrt{\mathbb{V}(\hat{\alpha} + \hat{\beta} x) + \mathbb{V}(\epsilon)}$
- We can construct confidence intervals and conduct hypothesis testing in the same manner as before


Looks and Politics in 2004 US Senate Elections
Margin of Victory (percentage points)
100
50
0
−50
−100
0.0
0.2
0.4
0.6
0.8
1.0
Competence (Proportions of participants)
alpha.hat = −38.4 (s.e. = 16.7)
beta.hat = 77.46 (s.e. = 16.66)

Statistical Inference with Multiple Regression

- Correlation does not imply causation
- Omitted variables $\Longrightarrow$ violation of exogeneity
- You can have multiple explanatory variables

$$Y_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_J X_{iJ} + \epsilon_i$$

- Interpretation of β_j: an increase in the outcome associated with one unit increase in X_{ij}
- Confidence intervals for $\hat{\beta}_j$, expected values, and predicted values can be constructed in the same manner
- Hypothesis testing for β_j, expected values, and predicted values, etc. can also conducted in the same manner

Electoral Costs of Iraq War Casualties

- Outcome: Change in Bush's vote share from 2000 to 2004
- Multiple regression from Karol and Miguel (*J. of Politics* 2007)
- The average number of casualties per 100,000 $= 3.39$

Variables	coef.	s.e.
Total Iraq deaths and wounded per 100,000	-0.0055	0.0023
Proportion active armed forces in 2000	0.43	0.26
Proportion veterans in 2000	-0.29	0.20
Change in unemployment, 9/2003 – 8/2004	-0.05	0.65
Change in Black pop. prop., 2000 – 2003	2.15	0.66
Change in White (non-Hispanic) pop. prop.	-0.35	0.60
Proportional change in total population	-0.12	0.16
Number of observations	51	
R^2 (coefficient of determination)	0.41	

R Regression Summary Output

```
Call:
lm(formula = log(gdppc) ~ ELF + oil, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-4.8182 -1.6462 -0.1049  1.3896  5.9684

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 24.883976   0.078750 315.986   <2e-16 ***
ELF         -0.023370   0.001653 -14.142   <2e-16 ***
oil         -0.003064   0.004527  -0.677    0.499
--
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 2.083 on 1998 degrees of free
Multiple R-squared: 0.0911, Adjusted R-squared: 0.0901
```

Going from Correlation to Causation

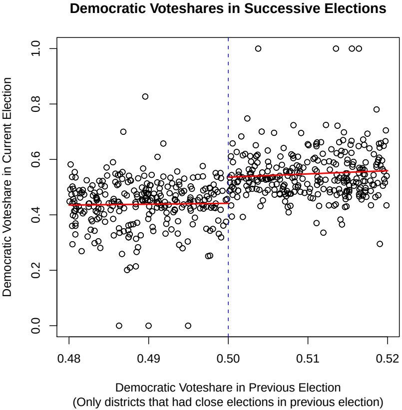
- Causal interpretation of linear regression: one unit increase in X_i changes Y_i by β holding other variables constant
- When can we give a causal interpretation?: X_i is randomized

- Randomized experiments are difficult to do
- Find natural experiments where "treatments" are assigned haphazardly
  - Vietnam draft lottery
  - Birthday
  - Weather
  - Assassination of political leaders

- Before and after design: Impact of increasing minimum wage
- Regression discontinuity design: take advantage of arbitrary cutoff points that naturally occur in the world
  - Admissions, Financial aid, Welfare eligibility, Class size, State and other administrative boundaries

Regression Discontinuity Design: Example

- Incumbent parties appear to be more likely to win elections but is it because they are incumbent?
- Use of close elections to estimate incumbency advantage



Lee et al. (2004) *Quarterly Journal of Economics*

Getting Ready for the Final Exam

- Final Exam:
  1. A take home exam
  2. 40% of the overall grade
  3. An open book exam (no outside materials) but no collaboration
  4. Posted on Jan 18 and due noon on Jan 23
  5. Another problem set (with conceptual and data analysis questions)
- Preparation: Review the materials you did not understand well. Do some practice problems including a practice final exam
- Review Session: Jan. 12, 3:30 – 4:30 pm, Bowl 2
- Office Hours:
  1. Kosuke: Jan 10, 2pm – 4pm (open door policy throughout the reading period)
  2. Carlos: Jan 13, 10am – noon (Corwin 023)
  3. Alex: Jan 17, noon - 2pm (Corwin 023)