# Lecture 29 (Ch. 9)

In Ch. 8, we tested k specific proportions in 1 pop, and in r pops

$H_0: \pi_1 = \pi_{01}, \pi_2 = \pi_{02}, \cdots, \pi_k = \pi_{0k}$ || $H_0$: pops are homog. w.r.t. categ.

$H_1$: At least 1 of these is wrong || $H_1$: not

chi-sqd dist. with df = k−1 || chi-sqd dist. with df = (k−1)(r−1)

We are skipping homogeneity

Now, how about k population means?

$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$ (Not $\mu_1 = \mu_{01}, \mu_2 = \mu_{02}, \cdots$)

$H_1$: At least 2 $\mu$'s are different.

The dist. turns out to be F with $(df_{num.}, df_{denom.})$. Recall test of model utility in regression.

The method is called 1-way (or single factor) ANOVA.

It deals with 1 continuous variable, y, whose mean is computed in k different levels of 1 categorical variable, x. ← "factor"

Example: Does knowledge of religion depend on religion?

THE PEW FORUM ON RELIGION & PUBLIC LIFE — A project of the Pew Research Center


TOPICS

ISSUES: Abortion, Church-State Law, Death Penalty, Education, Gay Marriage & Homosexuality, Government, Politics & Elections, Science & Bioethics, Social Welfare

BELIEFS & PRACTICES: Belief in God, Frequency of Prayer, Importance of Religion, Religious Attendance, Other Beliefs & Practices

RELIGIOUS AFFILIATION: Christian, Jewish, Muslim, Other Affiliations, Unaffiliated

DEMOGRAPHICS: Age, Education & Income, Gender, Geography, Race, Other Demographics

REGIONS: Americas, Asia & the Pacific, Europe, Middle East & North Africa, Sub-Saharan Africa

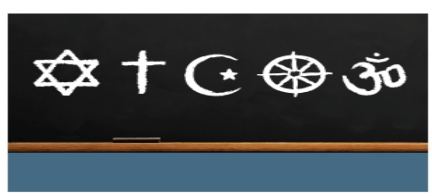
PUBLICATIONS: Analyses, Event Transcripts

pewforum.org › Topics › Beliefs & Practices


## U.S. Religious Knowledge Survey

POLL - September 28, 2010

### *Executive Summary*



Print Email Share AAA Text Size

Atheists and agnostics, Jews and Mormons are among the highest-scoring groups on a new survey of religious knowledge, outperforming evangelical Protestants, mainline Protestants and Catholics on questions about the core teachings, history and leading figures of major world religions.

On average, Americans correctly answer 16 of the 32 religious knowledge questions on the survey by the Pew Research Center's Forum on Religion & Public Life. Atheists and agnostics average 20.9 correct answers. Jews and Mormons do about as well, averaging 20.5 and 20.3 correct answers, respectively. Protestants as a whole average 16 correct answers; Catholics as a whole, 14.7. Atheists and agnostics, Jews and Mormons perform better than other groups on the survey even after controlling for differing levels of education.



**Atheists and Agnostics, Mormons and Jews Score Best on Religious Knowledge Survey**

*Average # of questions answered correctly out of 32*

| | |
|---|---|
| **Total** | **16.0** |
| Atheist/Agnostic | 20.9 |
| Jewish | 20.5 |
| Mormon | 20.3 |
| White evangelical Protestant | 17.6 |
| White Catholic | 16.0 |
| White mainline Protestant | 15.8 |
| Nothing in particular | 15.2 |
| Black Protestant | 13.4 |
| Hispanic Catholic | 11.6 |

test scores (out of 32)

Looks like Athiests know most!?

Even though we want to compare k means, it's not enough to look at sample means only.

SCIENTIFIC AMERICAN™

Winner of the 2011 National Magazine Award for General Excellence

Search ScientificAmerican.com

News & Features | Blogs | Multimedia | Education | Citizen Science | Topics

Home » Scientific American Magazine » December 2010

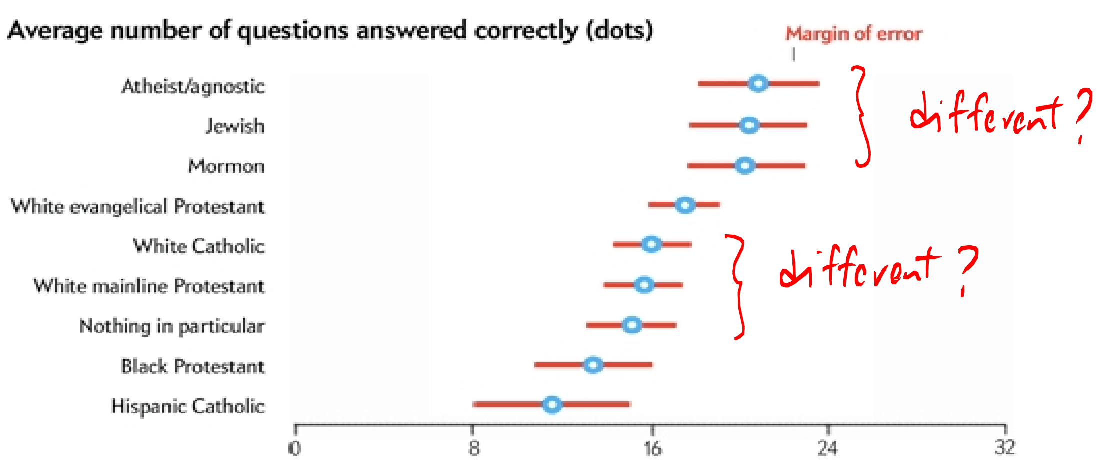
Advances | More Science


# The Science of "Disestimation": The Shortcomings of Opinion Polls

Why we shouldn't put our faith in opinion polls

By Charles Seife | December 14, 2010 | 19

Share Email Print




Moral: When testing means, std dev. matters!

Also note That This is just a generalization of The 2-sample/pop test (for comparing $\mu_1, \mu_2$) to The case of k populations.

## Example 9.1 (p. 410 – 411)

Vibration ($y$) for 5 brands ($x$) of bearing:

(or "speed" for 5 brands ($x$) of computers.)

Data:

| Brand 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 13.1 | ⋮ | ⋮ | ⋮ | ⋮ |
| 15.0 | | | | |
| 14.0 | | | | |
| ⋮ | | | | |
| 11.6 | | | | |
| $\bar{y}_1 = 13.68$ | $\bar{y}_2 = 15.97$ | 13.67 | 14.73 | 13.08 |
| $s_1 = 1.194$ | $s_2 = 1.167$ | | | |

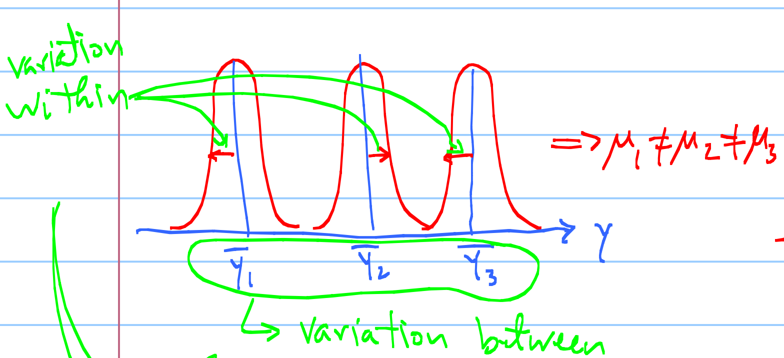
We want to know if the data provide evidence that the 5 bearings have different means (of vibration), $\mu_i$.

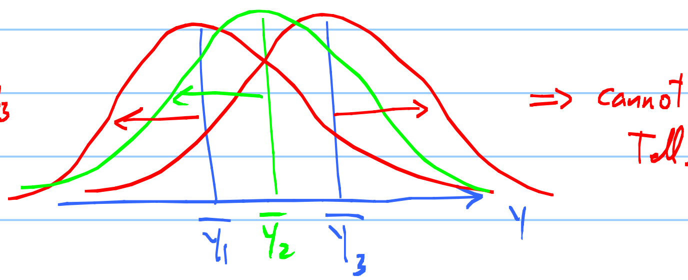
(i.e. Are the 5 computers different in their speed?)

$H_0$: $\mu_1 = \mu_2 = \cdots = \mu_5$ ← Note we are talking about 5 populations,

$H_1$: At least 2 $\mu$'s are diff.

The way ANOVA answers that question is by finding out how much of the total variation in $y$ is "within" each category, and how much is "between" the categories.




See math, next page.

Recall The decomposition of SST from regression.
Similarly,

k population.

sample size in $i^{th}$ pop.

$n = \sum_{i=1}^{k} n_i$

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{\bar{y}})^2$$

Grand mean $= \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}$

$j^{th}$ response in The $i^{th}$ pop.

Show!

$= \sum_{i=1}^{k} \left(\frac{n_i}{n}\right) \bar{Y}_i$

⋮

sample mean in $i^{th}$ pop.

Sample var. in $i^{th}$ pop.

$$SST = \sum_{i=1}^{k} n_i (\bar{Y}_i - \bar{\bar{Y}})^2 + \sum_{i=1}^{k} \overbrace{\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}^{= (n_i - 1) s_i^2}$$

$SS_{\text{between group}}$ = $SST_r$ ← Treatment = $SS_{\text{explained}}$

$SS_{\text{within group}}$ = $SSE$ = $SS_{\text{unexplained}}$

| | | | | | |
|---|---|---|---|---|---|
| SS : | Total | = | between | + | within |
| df : | $n-1$ | = | $(k-1)$ | + | $(n-k)$ |

$k$ = # of levels in 1 factor (predictor) = # of pops.

FYI Note:
linear regression: $n - 1 = k + [n - (k+1)]$

$k$ = # of $\beta$'s

different $k$

$\boxed{Q1}$: If all The sample means are all equal, Then $SS_{\text{between}}$ is

A) $\sum n_i (\bar{\bar{y}})^2$  B) $\sum n_i$  C) $\sum n_i (\bar{Y}_i)^2$  $\boxed{D)\ 0}$

$\bar{Y}_1 = \bar{Y}_2 = \cdots = \bar{Y}_k$, Then $\bar{\bar{Y}} = \bar{Y}_i$

Note if all the means are equal, then $F \approx 0$

Now, we can compare $SS_{between}$ and $SS_{within}$:

Theorem:

If $H_0$ = True, $F = \dfrac{SS_{between}/(k-1)}{SS_{within}/(n-k)} = \dfrac{MS_{between}}{MS_{within}}$

has an F-distribution with params. $df = (k-1, n-k)$

All we need is Table VIII to give us areas (p-values).

One assumption of this Theorem is that the y's in each of the k populations are normal, and that they all have the same variance, ie. $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$.

Use qq plots to test these assumptions. This assumption is called "homoscedasticity"!

Example 9.1 (p. 410 – 411)

$H_0$: $\mu_1 = \mu_2 = \cdots = \mu_5$

$H_1$: At least 2 $\mu$'s are diff.

y = Response = vibration

x = factor = brand type.

Data:

| | Brand 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| | 13.1 | ⋮ | ⋮ | ⋮ | ⋮ |
| | 15.0 | | | | |
| | 14.0 | | | | |
| | ⋮ | | | | |
| | 11.6 | | | | |
| | $n_1 = 6$ | $n_2 = 6$ | | | $n_5 = 6$ |
| | $\bar{y}_1 = 13.68$ | $\bar{y}_2 = 15.97$ | ... | | |
| | $s_1 = 1.194$ | $s_2 = 1.167$ | | | |

To reject $H_0$ in favor of $H_1$, $F_{obs}$ must be

A) small enough

$\boxed{\text{B) large enough}}$

Again, if $H_0$ is True, then $F \approx 0$. The bigger the F, the more evidence against $H_0$.

$$\bar{\bar{y}} = \sum_{i=1}^{5} \left(\frac{n_i}{n}\right) \bar{y}_i = \left(\frac{6}{30}\right)(13.68) + \cdots = 14.22$$

$$SSTr = \sum_{i=1}^{5} n_i \left(\bar{y}_i - \bar{\bar{y}}\right)^2 = 6\,(13.68 - 14.22)^2 + \cdots = 30.88$$

$$SSE = \sum_{i=1}^{5} \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_i\right)^2 = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots$$

$$= (6-1)(1.194)^2 + \cdots \simeq 22.83$$

$$F_{obs} = \frac{30.88/(5-1)}{22.83/(30-5)} = 8.45 \qquad df = (5-1,\ 30-5)$$

$$p\text{-value} = (\text{Table } \underline{VIII}) < .001$$

Conclusion:

∴ Reject $H_0$ ($\mu_1 = \mu_2 = \cdots$) in favor of $H_1$ (At least 2 $\mu$'s are different.) at $\alpha = .01$.

In "English":

∴ Sufficent evidence To reject "equal means" in favor of " at least 2 $\mu$'s are different." (which 2?)

∴ Bearing Type "has" an effect on vibration.
(At least one of The computers is faster than The others)

Section 9.3

We are skipping This, → Tukey's Test

Most software produce an ANOVA Table for keeping Track of all The relevant numbers, similar to regression. The structure is :

| Source | df | SS | Mean<br>MS | $F_{obs}$ | P-value |
|---|---|---|---|---|---|
| Between Group (factor) | $k-1$ | $SS_{between}$ | | | table VIII |
| | | from formula | | | |
| Within Group (error) | $n-k$ | $SS_{within}$ | | | |
| Total | $n-1$ | $SS_{Total}$ | | | |

In Lab. you will produce The ANOVA Table for The bearing example. You will find:

Response = vibration

Factor = type of bearing (5 levels)

| Source | df | SS | MS | F | P-value |
|---|---|---|---|---|---|
| factor | $5-1$ | 30.85 | 7.71 | 8.44 | .00018 |
| Error | $30-5$ | 22.84 | 0.91 | | |
| Total | $30-1$ | 53.7 | | | |

## Summary:

$\sigma$ unknown  $\sigma$ known

| | | | |
|---|---|---|---|
| z, t | $H_0: \mu = \mu_0$ | $H_1: \mu \square \mu_0$ | small/large sample |
| z | $H_0: \pi = \pi_0$ | $H_1: \pi \square \pi_0$ | large sample |
| z, t | $H_0: \mu_1 - \mu_2 = \Delta_0$ | $H_1: \mu_1 - \mu_2 \square \Delta_0$ | indep. or paired |

Chisqd $H_0: \pi_1 = \pi_{01}, \pi_2 = \pi_{02}, \cdots \pi_k = \pi_{0k}$, $H_1$: At least 1 is wrong

Chisqd $H_0$: homogeniety of r pops w.r.t. k categ. $H_1$: not

Equivalently: $H_0$: 2 categ. variables are independent $H_1$: not. skip!

F $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$  $H_1$: atleast 2 $\mu$'s are diff.

Note that the ANOVA F-test is a generalization of the 2-sample t-test to more than 2 populations.

---

I'm not assigning this as hw this quarter, because we did not learn how to do 1-way ANOVA in R. But if you are interested in doing it, let me know.

Do 1-way ANOVA on one of the 2 continous vars, and 1 of the categorical vars. in your data from hw-left1. If you cannot, explain why not!

Not assigned!