WELCOME TO THE 2019 SUMMER SEMINAR IN PHIL STAT

Deborah G. Mayo **Aris Spanos**

The Statistics Wars

2. Role of Probability: performance or probabilism? (Frequentist vs. Bayesian)

- Need new role:Probativism

Simple significance tests (Fisher)

p-value. …to test the conformity of the particular data under analysis with H_0 in some respect:

…we find a function $d(\mathbf{X})$ of the data, the **test statistic**, such that

- the larger the value of $d(\mathbf{X})$ the more inconsistent are the data with H_0;
- $d(\mathbf{X})$ has a known probability distribution when H_0 is true.

…the p-value corresponding to any $d(\mathbf{x})$ (or d_{0bs})

$$p = p(t) = Pr(d(\mathbf{X}) \geq d(\mathbf{x}); H_0)$$

(Mayo and Cox 2006, 81; d for t, x for y)

Testing Reasoning

- If even larger differences than d_{0bs} occur fairly frequently under H_0 (i.e., P-value is not small), there's scarcely evidence of incompatibility with H_0
- Small P-value indicates *some* underlying discrepancy from H_0 because **very probably you would have seen a less impressive** difference than d_{0bs} were H_0 true.

Popper ($\mu > \mu_1$) is corroborated (at level .975) because *it may be presented as a failed attempt to falsify it* statistically.

FEV(i) (p. 149): x is evidence against H_0 (evidence of a discrepancy from H_0), if and only if, were a correct description of the mechanism generating x, then, with **high probability, t**his would have resulted in a less discordant result than is exemplified by x (Mayo and Cox 2006, p. 82).

Popper ($\mu > \mu_1$) is corroborated (at level .975) because *it may be presented as a failed attempt to falsify it* statistically.

FEV(i) (p. 149): x is evidence against H_0 (evidence of a discrepancy from H_0), if and only if, were H_0 a correct description of the mechanism generating x, then, with **high probability, t**his would have resulted in a less discordant result than is exemplified by x (Mayo and Cox 2006, p. 82).
FEV/SEV (i) (p. 149): x is evidence against H_0 (evidence of a discrepancy from H_0), if and only if, the P-value $\Pr(d \geq d_0; H_0)$ probability is very low (equivalently, $\Pr(d < d_0; H_0) = 1 - \text{P-value}$ is very high.
Small

Here's your severity: there's a high probability the results would have gone (more) in a direction against what you're taking to infer H, if H is false.

Pr (Evidence in disagreement with your inference to C against Ho; Ho) would be **high**
You get inference that accords with C
Data indicate (or supply evidence for C)

and the warrant is this high probability

Variations on Test T+ (new ex starts p. 29)

Test (1)

1-sided normal testing

H_0: μ ≤ 150 vs. H_1: μ > 150 (Let σ = 10, n = 100)

let significance level α = .025

Reject H_0 whenever M ≥ 150 + 2σ/√n:

M ≥ 152

M is the sample mean, its value is M_0.

1SE = σ/√n = 1

P-value:

Pr(M ≥ 152; μ = 150)

Turn M into a Z statistics

Z = (152 – 150)/ SE

(observed - hypothesized)/SE

Z = 2

So want Pr (Z ≥ 2)

Look up Normal curve, answer is .025

(σ = 10, *n* = 100), 1SE = σ/√*n* = *10/10 =1*

So,

Pr(Test would have yielded a smaller difference from Ho, Ho)= .97

If didn’t, it yielded Z > 2

So there’s evidence of departure from Ho (i.e., $\mu > 150$)

And the .97 measures the well-testedness

FEV/SEV (i) (p. 149): x is evidence against H_0 (evidence of a discrepancy from H_0), if and only if the P-value

$Pr(d \geq d_0; H_0)$ is very low (equivalently, $Pr(d < d_0; H_0) = 1 – P\text{-value}$ is very high

Keeping M = 152 Consider different tests

Test 2

1-sided normal testing

H_0: μ ≤ 151 vs. H_1: μ > 151 (Let σ = 10, n = 100)

P-value associated with M = 152

Pr(M > 152; Ho: 151) = Pr(z > (152-151)/1) = Pr(Z > 1) on standard Normal Curve: .16

P-value = .16; 1 – p value = .84

Pr(Test would have yielded a smaller diff if Ho were true) = .84

some indication

Keeping M = 152 Consider different tests

Test 3

1-sided normal testing

H_0: μ ≤ 152 vs. H_1: μ > 152 (Let σ = 10, *n* = 100)

P-value associated with M = 152

Pr(M > 152; Ho: 152) = Pr(z > (152-152)/1) = Pr(Z > 0) on standard Normal Curve: .50

P-value = .5; 1 – p value = .5

Pr(Test would have yields a smaller diff if Ho were true) = .5

Lousy evidence

Go back to original test: 1-sided normal testing

H_0: μ ≤ 150 vs. H_1: μ > 150 (Let σ = 10, *n* = 100)

We don't want to infer merely there's some discrepancy from 150, we want to know the kind of magnitude information we saw

M = 152 is good evidence of μ > 150 SEV was .97

M = 152 is decent/some evidence of μ > 151 SEV was .84

M = 152 is lousy evidence of μ > 152 SEV was .5

Second third and 4th entries of p. 144.