Unit 10: Hypothesis Tests

Effect of sample size

Valesky vs. Brown: Both surveys below say 54% for V., so for CI, EV of sample % is .54. And for sig test: H_0: $p = 0.5$, H_a: $p > 0.5$.

- $n = 100$: For CI, SE $= \sqrt{.54(.46)/100} = .05$, so $54\% \pm 10\%$. For sig test, SE $= \sqrt{.5(.5)/100} = .05$, so $P(\% = .54) = P(z = (.54 - .5)/.05 = .8) = 21\%$
- $n = 1600$: For CI, SE $= \sqrt{.54(.46)/1600} = .0125$, so $54\% \pm 2.5\%$. For sig test, SE $= \sqrt{.5(.5)/1600} = .0125$, so $P(\% = .54) = P(z = (.54 - .5)/.0125 = 3.2) = .07\%$

Sgn test: Are x,y really related?

Regression line for data in sample approximates the regr line for population: $y = \alpha + \beta x$, where α, β are the intercept and slope for the population.

Are x,y for the population really related, i.e., is $\beta \neq 0$?

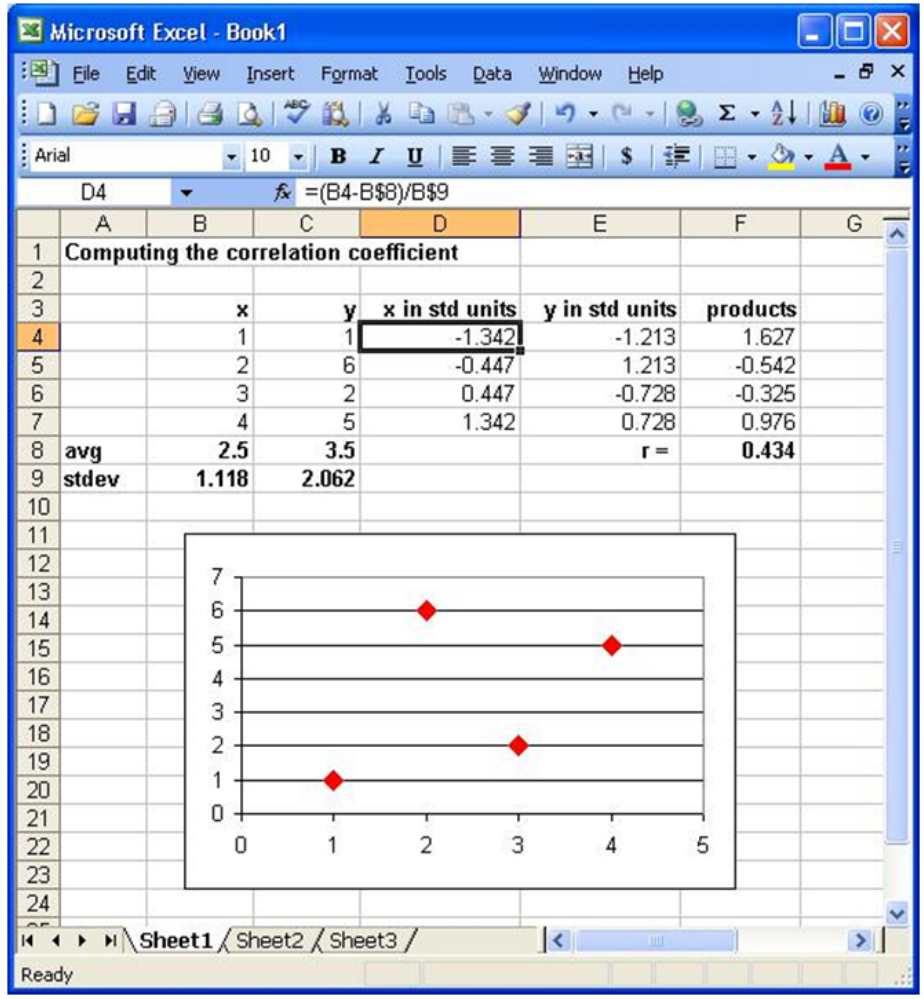
Sgn test:

- H_0: $\beta = 0$,
- $t = \sqrt{n-2} \cdot r / \sqrt{1-r^2}$,
- $df = n - 2$

Some statisticians, like our authors, disapprove of this test.
(Spreadsheet example: Unrelated points)

Example:

H_0: $\beta = 0$



$t = \sqrt{4-2}(.434)/\sqrt{1-(.434)^2} \approx .681$, $df = 4-2 = 2$

$P(t \geq .681 \text{ or } \leq -.681)$ is not less than 5%, so we fail to reject the null hypothesis: x and y are not related.