§6.3 Use and abuse of tests

- don't just say significant/not significant, report p-value.
- in large data sets, often find small statistically significant correlations, but not of much practical significance
- bad data sets not useful
- garden of forking paths.

§6.4 Power and inference

H_0 null hypotheses.
H_a alternate hypothesis

decision based on sample		Population: H_0 true	H_a true
	Reject H_0	Type I error α	correct decision
	Accept H_0	correct decision	Type II error. 1 − power

choose α: this means we reject H_0 if H_0 is true only α amount of the time, i.e. α is the probability we make a Type I error.

Q: what about type II: we accept (do not reject) H_0 even though H_a is true?
want this prob. to be as small as possible, say prob = 1 − power
i.e. want power of test to be as large as possible.

Example 6.17: water quality: test for lead parts per billion $N(\mu, \sigma)$ (assume $\sigma = 0.25$). EPA says limit is 15 ppb.

$H_0: \mu_0 = 15$
$H_a: \mu \neq 15$ (two sided)

$\alpha = 0.01$

sample: 15.84, 15.33, 15.58 $\qquad \bar{x} = \dfrac{15.84 + 15.33 + 15.58}{3} = 15.58$

test statistic: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \dfrac{15.58 - 15.00}{0.25/\sqrt{3}} = 4.02$