standard error

- standard error = standard deviation of a statistic
    - e.g., standard error of the mean
    - error = random variability due to random sampling ≠ mistake

- sampling distribution of the sample mean ($\bar{x}$; based on n samples)
    - e.g., body temperature, normal, population mean 37, population sd 0.5
        - consider mean of 100 samples; varies from sample to sample; what is its distribution?
    - standard deviation of $\bar{x}$ = <u>standard error</u> of the mean
        - $\sigma_{\bar{x}} = \sigma / \sqrt{n}$
        - here, $\sigma_{\bar{x}}$ = 0.5 / √100 = 0.5/10 = 0.05
    - if population is normal, then sampling distribution of mean is normal
    - central limit theorem: for n > 30, sampling distribution of the mean is approximately normally distributed, <u>even if the population is not normally distributed!</u>

    - what if we don't know the population standard deviation?
        - estimate it, with sample standard deviation s
        - then, estimate standard error: $s_{\bar{x}} = s/\sqrt{n}$
        - e.g., our 100 samples of body temperature has s = 0.45
            $s_{\bar{x}} = s/\sqrt{n}$ = 0.45 / 10 = 0.045

- t value of the sample mean
    t = (x - mu) / s_xbar
    - if n > 120, t is approximately normally distributed with mean 0, sd 1
        - use last line of appendix B
    - if n <= 120, t is not normally distributed
        - follows the t distribution with df = n - 1 "degrees of freedom"
        - t distribution is unimodal, symmetric, asymptotic; much like normal dist.
        - use appendix B
    - e.g., suppose we have 30 samples with mean 37.5 and standard deviation 0.52. if the population mean was 37, would this be in the top 5% of the sampling distribution? i.e., would this sample mean be unusually high?
        $s_{\bar{x}}$ = 0.52 / √30 = 0.0949
        t = (37.5 - 37) / 0.0949 = 5.27, df = 30 - 1 = 29
        - appendix B shows that for the top 5%, we need t >= 1.699; so yes, the mean 37 is in the top 5% of the sampling distribution
    - e.g., continuing above example, is the sample mean in the outer 20% of the sampling distribution of the mean? (i.e., is this unusually high or low?)
        - as before, $s_{\bar{x}}$ = 0.0949
        - as before, t = 5.27, df = 50 - 1 = 49
        - appendix B shows that we need |t| >= 1.311; so yes, the sample mean is in the outer 20% of the sampling distribution

significance testing

- thinking through a one sample t test
    - we have a population; we don't know its mean or standard deviation
    - we have a hypothesis about its mean, say H_0: $\mu = 20$; we want to test this
    - we collect n = 25 scores, and find the sample mean $\bar{x} = 22$ and sample standard deviation s = 2
    - how far is this sample mean from the hypothetical mean $\mu = 20$?
        - sample mean has standard error $s_{\bar{x}} = 2 / \sqrt{25} = 0.4$

        $t = (\bar{x} - \mu) / s_{\bar{x}} = (22 - 20) / 0.4 = 5$
    - is this sample mean (and its t statistic) consistent with the null hypothesis?
        - how large would the t statistic have to be in order to be in the outer 5% of the sampling distribution under the null hypothesis?
        - Appendix B: two-tailed, alpha = 0.05, df = 24: t = 2.064
        - we would need t < -2.064 or t > 2.064; and we do
    - so we reject the null hypothesis; it is probably not true that mu = 20

- worked example with mean shoe size, pp. 38-39; modified for two-tailed t test
    - one-sample t test
    - null hypothesis H_0: population of shoe sizes is normally distributed with mean $\mu = 9$
    - test this hypothesis in a two-tailed test with $\alpha = 0.05$
    - sample: 30 shoe sizes; mean = 10, sd = 2
    - if mean is 9, and sd is 2, then sampling distribution of the mean is normal with $\mu = 9$, $s_{\bar{x}} = 2 / \sqrt{30} = 0.3651$
    - t value of sample mean: t = (10 - 9) / 0.3651 = 3.7390
    - appendix B: with $\alpha = 0.05$, two-tailed, df = 29, we need |t| > 2.045
    - so yes, we reject the null hypothesis at the p = 0.05 level

- how small should p be, before we reject the null hypothesis?
    - α level, e.g., p = 0.05
    - type I error rate; false alarm; increases with alpha level
    - type II error rate; miss; harder to estimate; depends on alternatives; decreases with alpha level