Summary of statistical inference.

1) Hypothesis testing.

To test

$$H_0: \mu = \mu_0 \quad \text{against} \quad H_A: \mu \neq \mu_0,$$

we compute the z-score

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

If $|z| \geq \begin{cases} 1.96 \\ 2.33 \\ 2.576 \end{cases}$ then we reject H_0 and accept H_A at the $\begin{cases} 95\% \\ 98\% \\ 99\% \end{cases}$ level, respectively. Otherwise, we do not.

2) Confidence intervals.

A 95%/98%/99% confidence interval for a population mean is given by the interval

$$\left(\bar{x} - L\cdot\frac{s}{\sqrt{n}},\ \bar{x} + L\cdot\frac{s}{\sqrt{n}}\right),$$

where $L = 1.96/2.33/2.576$ respectively.

We use the same interval for a population proportion p, except in this case

$\bar{x} = \hat{p}$ (the sample proportion) and

$$s = \sqrt{\hat{p}(1-\hat{p})}.$$

3) In general,

"reject $H_0: \mu = \mu_0$ at the $p\%$ level"

is the same as

"μ_0 lies <u>outside</u> the $p\%$ confidence interval for μ."

New topic: <u>the binomial distribution.</u>

A <u>binomial</u> experiment is one with only two possible outcomes, often called a "success" and a "failure." Examples:

(a) flip a coin: we might call "heads" (or "tails") a "success."

(b) Shoot a free throw: a made shot is a success.

(c) Crossbreed pea plants: assuming only green or yellow plants can result, we might call "green" a success.

In general, let's write p for $P(\text{success})$.

E.g. for a fair coin, $p = 1/2$; for a free throw, p = shooter's free throw success rate.

NOW, suppose a binomial experiment is repeated n times. Two questions:

(1) What is the expected number of successes?

(2) Given an integer k with $0 \le k \le n$, what is the probability of exactly k successes out of the n trials?

ANSWERS:

(1) On a given trial, since $P(\text{success}) = p$, we would expect p successes.

So, if X_j denotes the number of successes on the j^{th} trial, then $E(X_j) = p$.

If X denotes the total number of successes out of n trials, then

$$X = X_1 + X_2 + \dots + X_n, \text{ so}$$

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \dots + X_n) \\ &= E(X_1) + E(X_2) + \dots + E(X_n) \\ &= \underbrace{p + p + \dots + p}_{n \text{ times}} = np. \end{aligned}$$

Example:

An 80% free throw shooter takes 15 shots in a game. The expected number of made shots is

$$np = 15 \cdot 0.8 = 12.$$

(2) We can think of it this way:

(a) There are $\binom{n}{k}$ of distributing the k successes among the n trials.

(b) For each of these ways, each of the k successes has probability p, and each of the $n-k$ failures has probability $1-p$. So the probability of the k successes and $n-k$ failures happening together is

$$p^k (1-p)^{n-k}.$$

(c) SO:

$$P(\text{exactly } k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k (1-p)^k.$$

Example: if an 80% free throw shooter takes 15 shots,

$$P(\text{exactly 12 made shots}) = \binom{15}{12} \cdot (0.8)^{12} \cdot (0.2)^3$$

$$\approx 0.250 = 25\%.$$

Thought experiment: without doing any computation, put these numbers (in the above free-throw context) in increasing order:

P(11 made), P(12 made), P(13 made).