Math 373 Lecture 12

CONVENTION. If q is a parameter for a population, the corresponding statistic for a sample is denoted by $\hat{q}$.

In a *binomial population*, the units are classified as successes or failures and p is the proportion or probability of success. To estimate the proportion p of successes in the population, select a random sample of an appropriate size n and measure the proportion $\hat{p}$ of successes in the sample.

The original population with probability p of success. The sample population with probability $\hat{p}$ of success. The *sampling* population is the set of the proportions $\hat{p}$ of the n-element samples. Different samples have different proportions $\hat{p}$. Taking the average and std. dev. of these sample proportions $\hat{p}$ gives the mean $\mu_{\hat{p}}$ and std. dev. $\sigma_{\hat{p}}$ of the *sampling distribution* of $\hat{p}$.

BINOMIAL PROPORTION THEOREM. Suppose n-element samples are taken from a binomial population with probabilities p and q of success and failure. Then

- the mean and std. dev. of the $\hat{p}$ are:

  $\mu_{\hat{p}} = p, \quad \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$ and
- the sampling distribution of $\hat{p}$ is approximately normal if the expected number of successes and failures in the sample exceeds 5, i.e., $np > 5$ and $nq > 5$.

■ A fair coin is tossed 100 times. Let $\hat{p}$ be the proportion of the tosses which are heads.

(a) Find the expected value of $\hat{p}$ and its std. dev.

We can regard tossing a coin 100 times as a sample of size $n = 100$ taken from the binomial population of all possible coin tosses with $p = .5$. Thus the theorem applies.

$$\mu_{\hat{p}} = p = .5, \quad \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{.5 \cdot .5}{100}} = \frac{.5}{10} = .05$$

(b) Is the distribution normal? Yes, $np = nq = 100 \times .5 > 5$.

(c) Within what limits do 95% of the proportions $\hat{p}$ lie?

Answer: $\mu \pm 2\sigma = .5 \pm 2(.05) = [.4, .6]$. Write in both forms.

(d) What is more likely: getting ≤ 4 heads in 10 tosses or getting ≤ 40 heads in 100 tosses?

THEOREM. Suppose x and y are independent random variables and a a constant. Then

$E(x+y) = E(x) + E(y)$,

$E(ax) = aE(x)$,

$\mathrm{Var}(x+y) = \mathrm{Var}(x) + \mathrm{Var}(y)$ ($\mathrm{Var}(x) = \sigma^2$ = the variance of x.)

$\mathrm{Var}(ax) = a^2 \mathrm{Var}(x)$,

$\sigma_{x+y} = \sqrt{\mathrm{Var}(x) + \mathrm{Var}(y)} = \sqrt{\sigma_x^2 + \sigma_y^2}$

$\sigma_{ax} = |a|\, \sigma_x$.

PROOF. This is from probability - Math 371.

COROLLARY. Suppose $x_1, x_2, \ldots, x_n$ are independent random variables with the same mean μ and std .dev. σ. Then

(a) The sum $x_1+x_2+\ldots+x_n$ has

mean $= n\mu$, variation $= n\sigma^2$, and std. dev. $= \sqrt{n}\,\sigma$

(b) (Central Limit Theorem)

The average $\bar{x} = (x_1+x_2+\ldots+x_n)/n$ has mean $\mu_{\bar{x}} = \mu$ and std. dev. $\sigma_{\bar{x}} = \mathrm{SE} = \frac{\sigma}{\sqrt{n}}$.

PROOF. (a)

mean $= E(x_1+x_2+\ldots+x_n) = E(x_1)+E(x_2)+\ldots+E(x_n)$

$= \mu+\mu+\ldots+\mu = n\mu$.

$\mathrm{Var}(x_1+x_2+\ldots+x_n) = \mathrm{Var}(x_1)+\mathrm{Var}(x_2)+\ldots+\mathrm{Var}(x_n)$.

$= \sigma^2+\sigma^2+\ldots+\sigma^2 = n\sigma^2$.

std. dev. $= \sqrt{n\sigma^2} = \sqrt{n}\,\sigma$.

(b) Let $S = x_1+x_2+\ldots+x_n$.

By (a), $E(S) = n\mu$ and $\sigma_S = \sqrt{n}\,\sigma$.

Since $\bar{x} = (x_1+x_2+\ldots+x_n)/n = S/n$,

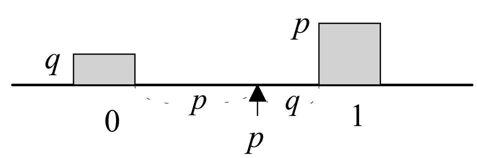
$\mu_{\bar{x}} = E(S/n) = (1/n)E(S) = (1/n)n\mu = \mu$.

$\sigma_{\bar{x}} = \sigma_{\frac{1}{n}S} = \frac{1}{n}\sigma_S = \frac{1}{n}\sqrt{n}\,\sigma = \frac{\sigma}{\sqrt{n}}$.

PROOF OF THE BINOMIAL SUCCESS THEOREM. Stated in Lecture 5.

One toss: Suppose you toss a bent coin once with probability p of heads. Let S be 1 if heads, 0 if tails. Thus S numerically encodes the success (heads) or failure of the one toss.

What is the mean and std. dev. of S?



mean $= E(S) = 0(q) + 1(p) = p$

$$\sigma_S = \sqrt{(0-p)^2(q) + (1-p)^2(p)}$$

$$= \sqrt{p^2 q + q^2 p} = \sqrt{pq(p+q)} = \sqrt{pq} .$$

n-tosses: Suppose this bent coin is tossed n times and x is the number of heads. Let $s_i = 1$ if the i^{th} toss is heads, 0 if not. Then the s_i's are independent and have mean $\mu = p$ and std. dev. $\sigma = \sqrt{pq}$.

$0+1+0+1+1 = 3$ = the number of 1's in the sum. Thus $s_1+s_2+\ldots+s_n$ = the number of 1's = the # of heads = x.

By Corollary (a) the expected number of heads is $\mu = E(x) = E(s_1+s_2+\ldots+s_n) = np$ with std. dev. $\sigma = \sqrt{n}\,\sqrt{pq} = \sqrt{npq}$.

PROOF OF THE BINOMIAL PROPORTION THEOREM. The proportion $\hat{p}$ of successes $= x/n = (s_1+s_2+\ldots+s_n)/n$ which is the average of the s_i's. Each s_i has mean $\mu = p$ and std. dev. $\sigma = \sqrt{pq}$. By Corollary (b), $\hat{p}$ has mean $\mu_{\hat{p}} = p$ and std. dev. $\sigma_{\hat{p}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{pq}}{\sqrt{n}} = \sqrt{\frac{pq}{n}}$.