Repeated Trials and Sampling: Summary

Binomial Probability Formula

$$P(k \text{ successes in } n \text{ trials}) = \binom{n}{k} p^k q^{n-k} \qquad \text{for independent trials with}$$

$$p = \text{probability of success on each trial,}$$
$$q = 1 - p = \text{probability of failure on each trial.}$$

For fixed n, as k varies from 0 to n, these probabilities define the *binomial* (n, p) *distribution* on $\{0, 1, \ldots, n\}$. That the probabilities add to 1 amounts to the

Binomial Theorem: $(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$

Here,
$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$$
= binomial coefficient called n *choose* k

= number of ways to pick k places out of n

= number of subsets of k of a set of n

= number in row n, column k of Pascal's triangle

Note: $\binom{n}{n} = \binom{n}{0} = 1$

Recursion Formula for Pascal's Triangle

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \qquad (\text{for } 0 < k < n, \quad n = 1, 2, \ldots)$$

Symmetry of Pascal's Triangle

$$\binom{n}{k} = \binom{n}{n-k}$$

Consecutive Ratios in Pascal's Triangle

$$\binom{n}{k} \Big/ \binom{n}{k-1} = \frac{n-k+1}{k}$$

Consecutive Ratios in the Binomial (n, p) Distribution

$$R(k) = \frac{P(k)}{P(k-1)} = \frac{(n-k+1)}{k}\frac{p}{q}$$

Mode of Binomial (n, p) Distribution: m = most likely value = $\text{int}(np+p)$

Normal Approximation to the Binomial Distribution

$$P(k) \approx \frac{1}{\sigma}\phi\left(\frac{k-\mu}{\sigma}\right)$$

where $\mu = np$ is the *mean*,

$\sigma = \sqrt{npq}$ is the *standard deviation*,

$z = (k-\mu)/\sigma$ is k in *standard units*,

$\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$ is the *standard normal density function.*

$$P(a \text{ to } b) \approx \Phi\left(\frac{b+\frac{1}{2}-\mu}{\sigma}\right) - \Phi\left(\frac{a-\frac{1}{2}-\mu}{\sigma}\right)$$

where $\Phi(z) = \int_{-\infty}^{z}\phi(x)dx$ is the *standard normal c.d.f.*

This approximation should be used only if $\sigma \geq 3$. The larger σ, the better.

$$\Phi(-z) = 1 - \Phi(z)$$
$$\Phi(a,b) = \Phi(b) - \Phi(a)$$
$$\Phi(-b,b) = 2\Phi(b) - 1$$
$$P(\mu-\sigma \text{ to } \mu+\sigma \text{ success in } n \text{ trials}) \approx \Phi(-1,1) \approx 68\%$$
$$P(\mu-2\sigma \text{ to } \mu+2\sigma \text{ success in } n \text{ trials}) \approx \Phi(-2,2) \approx 95\%$$
$$P(\mu-3\sigma \text{ to } \mu+3\sigma \text{ success in } n \text{ trials}) \approx \Phi(-3,3) \approx 99.7\%$$

Square Root Law for Independent Trials: The deviation from the expected number of successes np will most likely be a small multiple of $\sigma = \sqrt{npq} \leq \frac{1}{2}\sqrt{n}$.

$$P(p - \tfrac{1}{\sqrt{n}} \leq \text{sample proportion} \leq p + \tfrac{1}{\sqrt{n}}) \geq 95\% \qquad \text{for large } n.$$

Poisson Approximation to the Binomial Distribution

If p is close to zero

$$P(k) \approx e^{-\mu}\mu^k/k! \qquad \text{where } \mu = np$$

Random Sampling: See box on page 125.