Bernoulli distribution

The simplest non-uniform distribution

\[ f(x) = P(X = x) = \begin{cases} 
  p & \text{if } x = 1 \\
  1 - p & \text{if } x = 0 
\end{cases} \]

Jacob Bernoulli (1654-1705)
Swiss mathematician (Basel)

- Law of large numbers
- Mathematical constant \( e = 2.718 \ldots \)
Bernoulli distribution

\[ f(x) = P(X = x) = \begin{cases} 
  p & \text{if } x = 1 \\
  1 - p & \text{if } x = 0 
\end{cases} \]

\[ E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = 0(1 - p) + 1(p) = p \]

\[ \text{Var}(X) = E(X^2) - (EX)^2 = [0^2(1 - p) + 1^2(p)] - p^2 = p - p^2 = p(1 - p) \]
Refresher: Binomial Coefficients

\[
\binom{n}{k} = C_k^n = \frac{n!}{k!(n-k)!}, \text{ called } n \text{ choose } k
\]

\[
\binom{10}{3} = C_{10}^3 = \frac{10!}{3!7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{3 \cdot 2 \cdot 1 \cdot 7!} = 120
\]

Number of ways to choose \( k \) objects out of \( n \) 
without replacement and where the order does not matter. 
Called binomial coefficients because of the binomial formula

\[
(p + q)^n = (p + q) \times (p + q) \times \cdots \times (p + q) = \sum_{x=0}^{n} C^n_x p^x q^{n-x}
\]
Binomial Distribution

• Binomially-distributed random variable $X$ equals sum (number of successes) of $n$ independent Bernoulli trials

• The probability mass function is:

$$f(x) = C^n_x p^x (1 - p)^{n-x} \quad \text{for } x = 0, 1, \ldots, n \quad (3-7)$$

• Based on the binomial expansion:

$$f = (p + q)^n = \sum_{x=0}^{n} C^n_x p^x q^{n-x}$$

Sec 3.6 Binomial Distribution
Binomial Mean and Variance

X is a binomial random variable with parameters $p$ and $n$.

Mean:
$\mu = E(X) = np$

Variance:
$\sigma^2 = V(X) = np(1-p)$

Standard deviation:
$\sigma = \sqrt{np(1-p)}$
Matlab exercise: Binomial distribution

• Generate a sample of size 100,000 for binomially-distributed random variable X with n=100, p=0.2

• Tip: generate n Bernoulli random variables and use sum to add them up

• Plot the approximation to the Probability Mass Function based on this sample

• Calculate the mean and variance of this sample and compare it to theoretical calculations: 
  \[ E[X] = n \cdot p \] and 
  \[ V[X] = n \cdot p \cdot (1-p) \]
Matlab template: Binomial distribution

- n=100; p=0.2;
- Stats=100000;
- r1=rand(Stats,n) ?? < or > ?? p;
- r2=sum(r1, ?? 1 rows or 2 columns ?? );
- mean(r2)
- var(r2)
- [a,b]=hist(r2, 0:n);
- p_b=??./sum(??);
- figure; stem(??,p_b);
- figure; semilogy(??,p_b,'ko-')
Matlab exercise: Binomial distribution

- n=100; p=0.2;
- Stats=100000;
- r1=rand(Stats,n)<p;
- r2=sum(r1,2);
- mean(r2)
- var(r2)
- [a,b]=hist(r2, 0:n);
- p_b=a./sum(a);
- figure; stem(b,p_b);
- figure; semilogy(b,p_b,'ko-')