A random variable $X$ has the same probability for integer numbers $x = 1:10$

What is its skewness?

A. 0.5  
B. 1  
C. 0  
D. 0.1

Get your i-clickers
An example of the uniform distribution

Cycle threshold (Ct) value in COVID-19 infection
What is the Ct value of a PCR test?

\[ Ct = \text{const} - \log_2(\text{viral DNA concentration}) \]
Why $C_t$ distribution should it be uniform?

$C_t = 15 - 25$

Random samples from test

$\text{PMF}(C_t)$

$\{15, 40\}$

$10^8$

$10^3$

5 days

detectable at $C_t \geq 40$

time
Examples of uniform distribution: Ct value of PCR test of a virus

![Graph showing distribution of cycle threshold (CT) values.](image)

**Figure 3** Distribution of cycle threshold (CT) values. The total number of specimens with indicated CT values for Target 1 and 2 are plotted. The estimated limit of detection for (A) Target 1 and (B) Target 2 are indicated by vertical dotted lines. Horizontal dotted lines encompass specimens with CT values less than 3x the LoD for which sensitivity of detection may be less than 100%. This included 19/1,180 (1.6%) reported CT values for Target 1 and 81/1,211 (6.7%) reported CT values for Target 2. Specimens with Target 1 or 2 reported as “not detected” are denoted as a CT value of “0.”

Distribution of SARS-CoV-2 PCR Cycle Threshold Values Provide Practical Insight Into Overall and Target-Specific Sensitivity Among Symptomatic Patients

Why should we care?

Non-mandatory tests in Israel

- High Ct value means we identified the infected individual early, hopefully before transmission to others

Mandatory tests at UIUC 2021

- When testing is mandatory, and people are tested frequently – Ct value is skewed towards high values
Bernoulli distribution

The simplest non-uniform distribution

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

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

- Law of large numbers
- Mathematical constant \( e = 2.718... \)
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

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

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

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

\[
(p + q)^n = (p + q)(p + q)\ldots(p + q) = \sum_{x=0}^{n} C^n_x p^x q^{n-x}
\]
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:

$$f = (p + q)^n = \sum_{x=0}^{n} C^n_x p^x q^{n-x}$$
Binomial variance and standard deviation

Let $X$ be a binomial random variable with parameters $p$ and $n$

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 p=0.2 and n=100

• 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 * p and V[X]=n*p*(1-p)
Matlab exercise: Binomial distribution

- Stats=100000; n=100; p=0.005;
- r1=rand(Stats,n)<p;
- r2=sum(r1,2);
- disp('Sample mean'); disp(mean(r2));
- disp('Expected mean:'); disp(n.*p);
- disp('Sample variance'); disp(var(r2));
- disp('Expected variance:'); disp(n.*p.*(1-p));
- [a,b]=hist(r2, 0:1:n);
- p_b=a./sum(a);
- figure; stem(b,p_b);
- figure; semilogy(b,p_b,'ko-')