Office hours

Gonsalves, Novia Francis, ngons2@illinois.edu

Jayne, Dalton, djayne2@illinois.edu

Office hours:
Everitt 1103
every Wednesday from
2:30-4:00pm

Tentative schedule of exams

BIOE 310 (Fa26)

Exams | Staff | Students | Log

Exams

Name	Labels	Visibility	Duration	Start	End	Self-reserve
Exam 1	—	Visible	50min	2026-09-29 00:01:00 (CDT)	2026-10-01 23:59:00 (CDT)	2026-09-17 00:01:00 (CDT)
Exam 2	—	Visible	50min	2026-11-03 00:01:00 (CST)	2026-11-05 23:59:00 (CST)	2026-10-22 00:01:00 (CDT)
Final Exam	—	Visible	1h 50min	2026-12-10 00:01:00 (CST)	2026-12-15 23:59:00 (CST)	2026-10-19 00:01:00 (CDT)

Summary: Parameters of a Probability Distribution

- Probability Mass Function (PMF): $f(x)=\text{Prob}(X=x)$
- Cumulative Distribution Function (CDF): $F(x)=\text{Prob}(X\leq x)$
- Complementary Cumulative Distribution Function (CCDF): $F_>(x)=\text{Prob}(X>x)$
- The mean, $\mu=E[X]$, is a measure of the center of mass of a random variable
- The variance, $V(X)=E[(X-\mu)^2]$, is a measure of the dispersion of a random variable around its mean
- The standard deviation, $\sigma=[V(X)]^{1/2}$, is another measure of the dispersion around mean. Has the same units as X
- The skewness, $\gamma_1=E[(X-\mu)^3/\sigma^3]$, a measure of asymmetry around mean
- The geometric mean, $\exp(E[\log X])$ is useful for very broad distributions

A gallery of useful
discrete probability distributions

Discrete Uniform Distribution

- Simplest discrete distribution.
- The random variable X assumes only a finite number of values, each with equal probability.
- A random variable X has a discrete uniform distribution if each of the n values in its range, say $x_1, x_2, \ldots, x_n$, has equal probability.

$f(x_i) = 1/n$

Uniform Distribution of Consecutive Integers

- Let X be a discrete uniform random variable all integers from a to b (inclusive). There are $b - a + 1$ integers. Therefore each one gets:

  $$f(x) = 1/(b-a+1)$$

- Its measures are:

  $$\mu = E(x) = (b+a)/2$$

  $$\sigma^2 = V(x) = [(b-a+1)^2 - 1]/12$$

Note that the mean is the midpoint of a & b.

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

What is the behavior of its Probability Mass Function (PMF): $P(X=x)$?

A. does not change with x=1:10
B. linearly increases with x=1:10
C. linearly decreases with x=1:10
D. is a quadratic function of x=1:10

Get your i-clickers

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

What is the behavior of its Cumulative Distribution Function (CDF): $P(X \le x)$?

A. does not change with $x=1{:}10$
B. $\boxed{\text{linearly increases with } x=1{:}10}$
C. linearly decreases with $x=1{:}10$
D. is a quadratic function of $x=1{:}10$

Get your i-clickers

A random variable X has the same probability for integer numbers

x =1:10

What is its mean value?

A. 0.5
B. $\boxed{5.5}$
C. 5
D. 0.1

Get your i-clickers

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

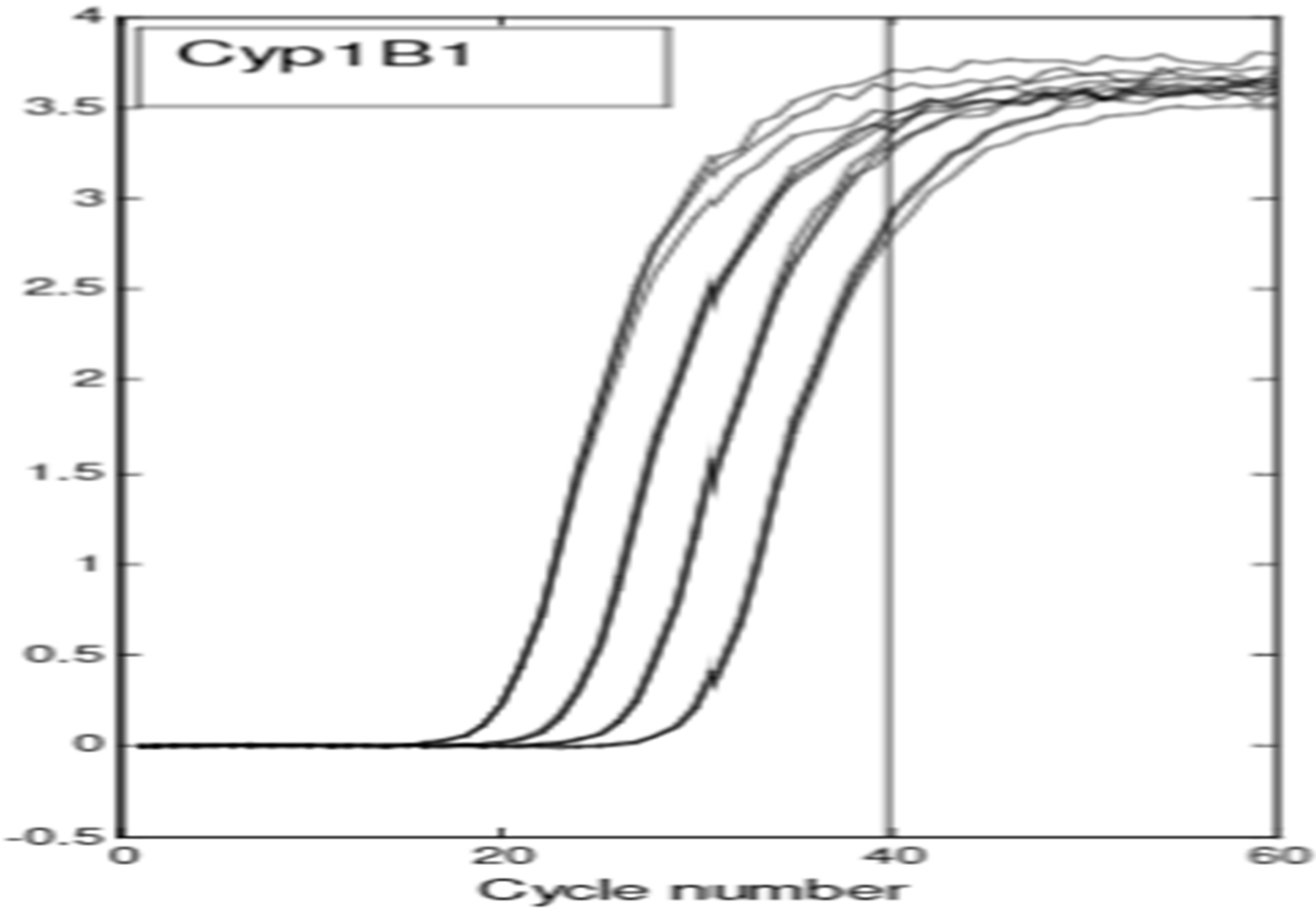
Get your i-clickers

An example of the uniform distribution

Cycle threshold (Ct) value in COVID-19 infection

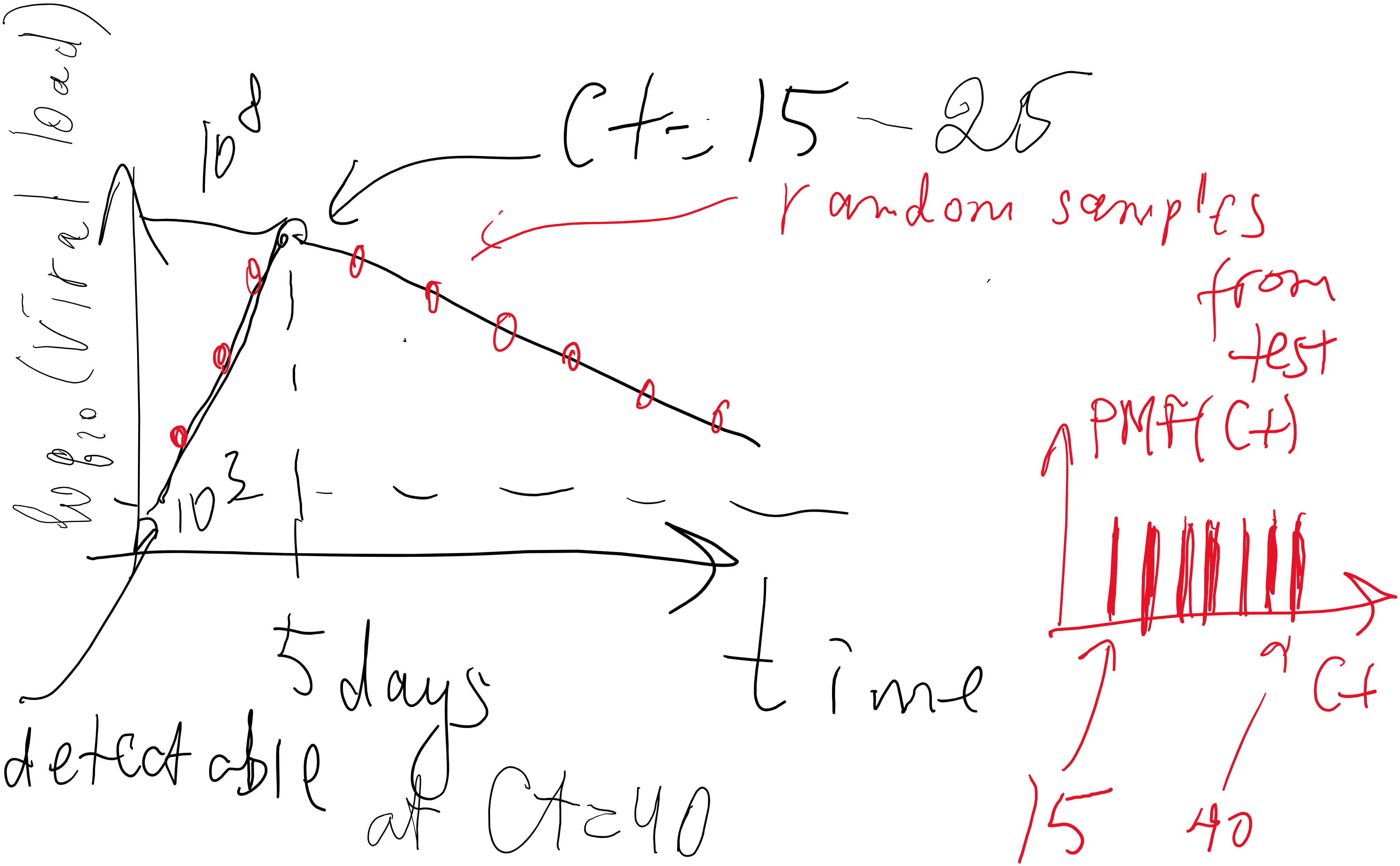
What is the Ct value of a PCR test?

Ct = const – log2(viral DNA concentration)



Why Ct distribution should be uniform?

Why Ct distribution should be uniform?



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

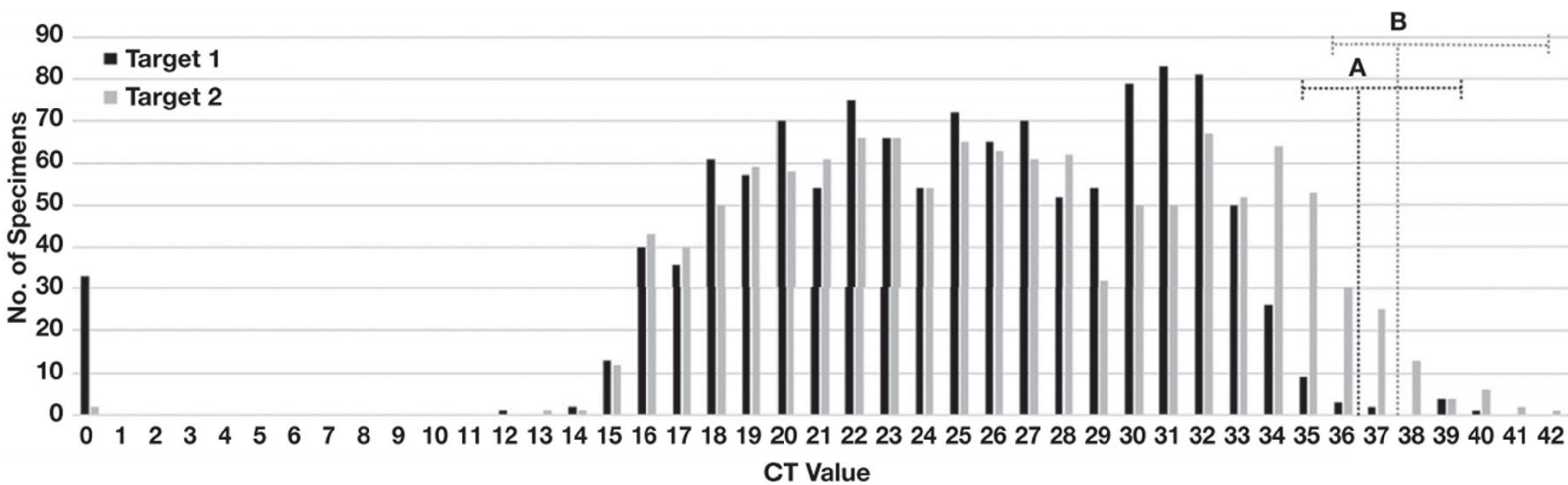


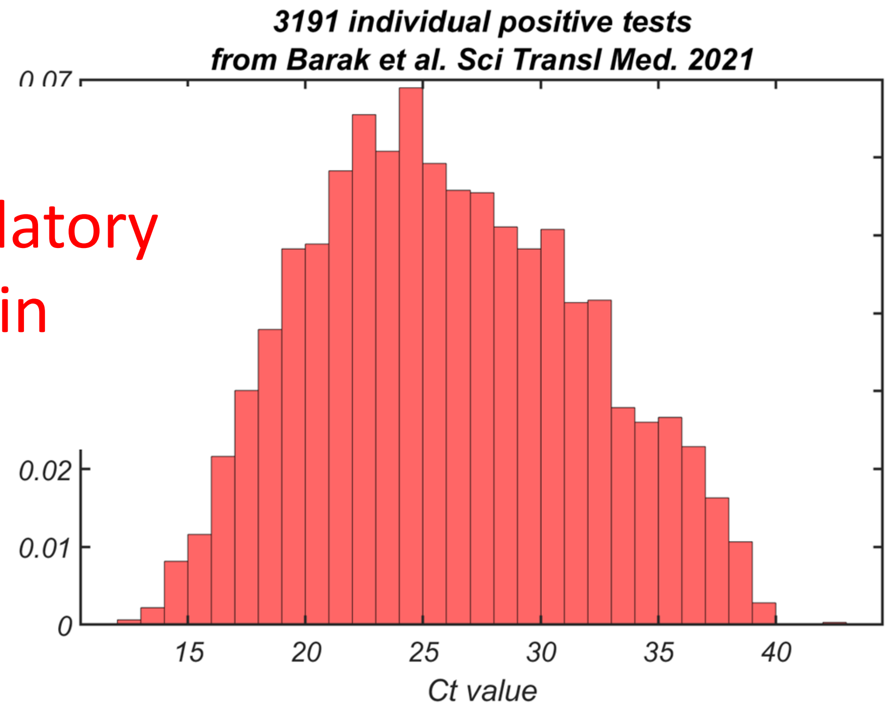
Figure 3 Distribution of cycle threshold (CT) values. The total number of specimens with indicated CT values for Target 1 and 2 are plotted. The estimated limit of detection for (A) Target 1 and (B) Target 2 are indicated by vertical dotted lines. Horizontal dotted lines encompass specimens with CT values less than 3× 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

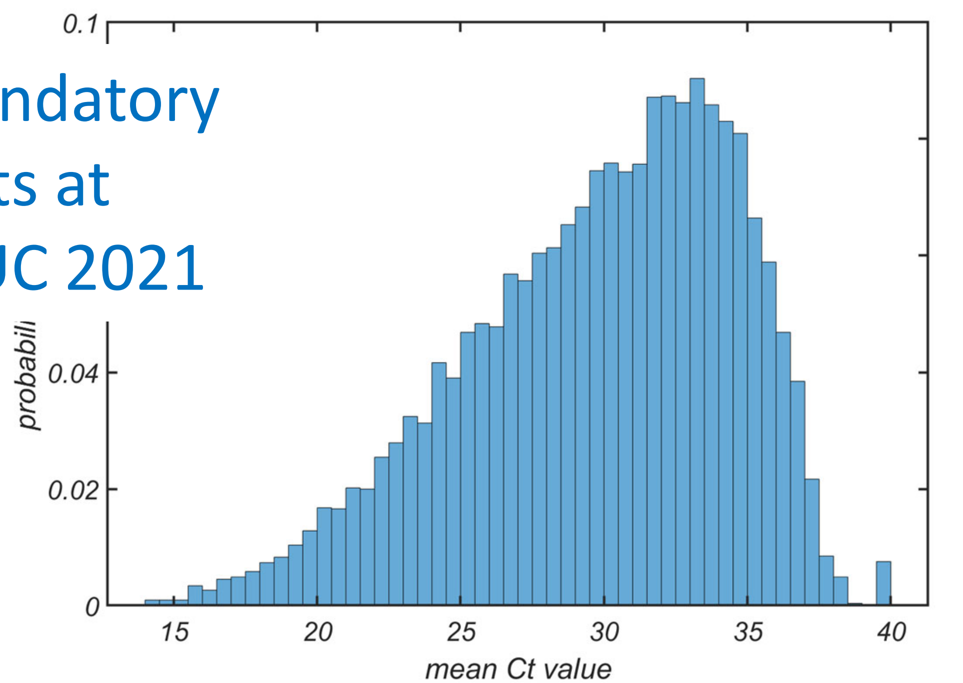
Blake W Buchan, PhD, Jessica S Hoff, PhD, Cameron G Gmehlin, Adriana Perez, Matthew L Faron, PhD, L Silvia Munoz-Price, MD, PhD, Nathan A Ledeboer, PhD *American Journal of Clinical Pathology*, Volume 154, Issue 4, 1 October 2020, https://academic.oup.com/ajcp/article/154/4/479/5873820

Why should we care?

Non-mandatory tests in Israel



Mandatory tests at UIUC 2021



- High Ct value means we identified the infected individual early, hopefully before transmission to others
- When testing is mandatory, and people are tested frequently – the mean Ct value is shifted towards higher values

Matlab exercise: Uniform distribution

- Generate a sample of size 100,000 for uniform random variable X taking values 1,2,3,…10
- Plot the approximation to the probability mass function based on this sample
- Calculate mean and variance of this sample and compare it to infinite sample predictions: $E[X]=(a+b)/2$ and $V[X]=((a-b+1)^2-1)/12$

Matlab template: Uniform distribution

- b=10; a=1; % b= upper bound; a= lower bound (inclusive)'
- Stats=100000; % sample size to generate
- r1=rand(Stats,1);
- r2=floor(??*r1)+??;
- mean(r2)
- var(r2)
- std(r2)
- [hy,hx]=hist(r2, 1:10); % hist generates histogram in bins 1,2,3...,10
- % hy - number of counts in each bin; hx - coordinates of bins
- p_f=hy./??; % normalize counts to add up to 1
- figure; plot(??,p_f, 'ko-'); ylim([0, max(p_f)+0.01]); % plot the PMF

Matlab exercise: Uniform distribution

- **b=10; a=1; % b= upper bound; a= lower bound (inclusive)'**
- **Stats=100000; % sample size to generate**
- **r1=rand(Stats,1);**
- **r2=floor(b*r1)+a;**
- **mean(r2)**
- **var(r2)**
- **std(r2)**
- **[hy,hx]=hist(r2, 1:10); % hist generates histogram in bins 1,2,3...,10**
- **% hy - number of counts in each bin; hx - coordinates of bins**
- **p_f=hy./sum(hy); % normalize counts to add up to 1**
- **figure; plot(hx,p_f, 'ko-'); ylim([0, max(p_f)+0.01]); % plot the PMF**

Bernoulli distribution

The simplest non-uniform distribution

p – probability of success (1)

1-p – probability of failure (0)

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

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

- Law of large numbers
- Mathematical constant e=2.718…

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

$$(p+q)^n = (p+q) \times (p+q) \ldots \times (p+q) = \sum_{x=0}^{n} C_x^n p^x q^{n-x}$$

Binomial Distribution

- Binomially-distributed random variable X equals sum (number of successes) of n independent Bernoulli trials
- The probability mass function is:

$q = 1 - p$

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

- Based on the binomial expansion:

$$1 = (p+q)^n = \sum_{x=0}^{n} C_x^n p^x q^{n-x}$$

Binomial Mean

X is a binomial random variable
with parameters *p* and *n*

Mean:

$\mu = E(X) = np$

$$\mu = \sum x C_x^n p^x q^{n-x} = p \frac{\partial}{\partial p} \sum C_x^n p^x q^{n-x} =$$

$$= p \frac{\partial}{\partial p} (p+q)^n = np$$

$$E(X(X-1)) =$$

$$= \sum x(x-1) C^n_x p^x q^{n-x}$$

$$= p^2 \frac{\partial^2}{\partial p^2} \sum C^n_x p^x q^{n-x} =$$

$$= p^2 \frac{\partial^2}{\partial p^2} (p+q)^n \Big|_{q=1-p} = n(n-1)p^2$$

$$E(X^2) = E(X(X-1)) + E(X) =$$

$$= n^2p^2 - np^2 + np = n^2p^2 + np(1-p)$$

$$V(X) = E(X^2) - E(X)^2 = n^2p^2 + np(1-p) - (np)^2$$

$$= \boxed{np(1-p)}$$

Binomial mean, variance and standard deviation

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

- Mean:

$\mu = np$

- Variance:

$\sigma^2 = V(X) = np(1\text{-}p)$

- Standard deviation:

$\sigma = \sqrt{np(1-p)}$

- Standard deviation to mean ratio

$$\sigma/\mu = \sqrt{np(1-p)}/np = \frac{\sqrt{(1-p)/p}}{\sqrt{n}}$$

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*p and V[X]=n*p*(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-')**