

# ST2334 - Special Probability Distributions

4.1 - Discrete Uniform Distribution

4.2 - Bernoulli and Binomial Distributions

4.3 - Negative Binomial Distribution

4.4 - Poisson Distribution

4.5 - Poisson Approximation to the Binomial Distribution

4.6 - Continuous Uniform Distribution

4.7 - Exponential Distribution

4.8 - Normal Distribution

4.9 - Normal Approximation to the Binomial Distribution

## 4.1 - Discrete Uniform Distribution

- Definition 4.1: Random variable  $X$  assumes the values  $x_1, x_2, \dots, x_k$ ,

with equal probability

→  $X$  has a discrete uniform distribution and p.f. is given by:

$$f_X(x) = \begin{cases} \frac{1}{k} & x = x_1, x_2, \dots, x_k \\ 0 & \text{otherwise} \end{cases}$$

- Theorem 4.1: Mean and variance of the discrete uniform distribution :

$$\mu = E(X) = \sum_{\text{all } x} x f_X(x) = \sum_{i=1}^k x_i \cdot \frac{1}{k} = \frac{1}{k} \sum_{i=1}^k x_i$$

$$\text{or } \sigma^2 = V(X) = \sum_{\text{all } x} (x - \mu)^2 f_X(x) = \frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2$$

$$\sigma^2 = E(X^2) - \mu^2 = \frac{1}{k} \left( \sum_{i=1}^k x_i^2 \right) - \mu^2$$

## 4.2 - Bernoulli and Binomial Distributions

### 4.2.1 - Bernoulli Distributions → special case of binomial distribution ( $n=1$ )

wrote 1 and 0

- Bernoulli experiment : a random experiment with only two possible outcomes

- Definition 4.2 : A random variable  $X$  has a Bernoulli distribution if p.f. of  $X$  is :

$$f_X(x) = p^x(1-p)^{1-x}, \quad x=0,1;$$

where parameter  $p$  satisfies  $0 < p < 1$ .

$$f_X(x) = 0 \text{ for other } X \text{ values}$$

-  $(1-p)$  is often denoted as  $q$

-  $\Pr(X=1) = p$  and  $\Pr(X=0) = 1-p = q$

- Theorem 4.2 : Mean and variance of a Bernoulli distribution :

$$\mu = E(X) = p$$

$$\sigma^2 = V(X) = p(1-p) = pq$$

### 4.2.2 - Binomial Distributions

i.e.  $X \sim B(n,p)$

- Definition 4.3 : A random variable  $X$  has a binomial distribution

with two parameters  $n$  and  $p$ , if p.f. of  $X$  is :

$$\Pr(X=x) = f_X(x) = \binom{n}{x} p^x(1-p)^{n-x} = \binom{n}{x} p^x q^{n-x},$$

for  $x=0,1,\dots,n$  where  $p$  satisfies  $0 < p < 1$ ,  $q = 1-p$ ,

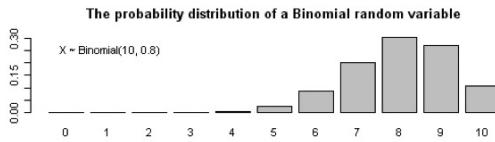
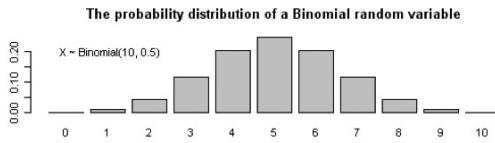
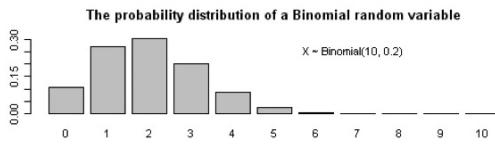
and  $n$  ranges over the positive integers

-  $X$  is the number of successes that occur in  $n$  independent Bernoulli trials

- Theorem 4.3 : Mean and variance of a binomial distribution : for  $X \sim B(n,p)$

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

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



- Conditions for a binomial experiment :

- ① Only consists of  $n$  repeated Bernoulli trials
- ② Only two possible outcomes : success and failure in each trial
- ③  $\Pr(\text{success}) = p$  is the same constant in each trial
- ④ Trials are independent

- The RV  $X$  is the number of successes among the  $n$  trials in a binomial experiment

$$\rightarrow X \sim B(n, p)$$

## 4.3 - Negative Binomial Distribution

- Let  $X$  be a random variable represents the number of trials to produce the  $k$  successes in a sequence of independent Bernoulli trials  $\xrightarrow{X \sim NB(k, p)}$
- $X$  follows a negative binomial distribution with parameters  $k$  and  $p$
- p.f. of  $X$ :  $\Pr(X = x) = f_X(x) = \binom{x-1}{k-1} p^k q^{x-k}$   
for  $x = k, k+1, k+2, \dots$

- Mean and variance of a negative binomial distribution:

$$E(X) = \frac{k}{p}$$

$$\text{Var}(X) = \frac{(1-p)k}{p^2}$$

$t$ , may be of any length e.g. minute, day, week, month, year

## 4.4 - Poisson Distribution

- Poisson experiments: experiments yielding numerical values of a random variable  $X$ ,  
the number of successes occurring during a given time interval  
or in a specified region
- Properties of Poisson experiments:
  - ① Number of successes occurring in one time interval or specified region are independent of those occurring in any other disjoint time interval or region of space
  - ② Probability of a single success occurring during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of successes occurring outside this time interval or region
  - ③ Probability of more than one success occurring in such a short time interval or falling in such a small region is negligible
- Definition 4.4: Poisson random variable: number of successes  $X$  in a Poisson experiment

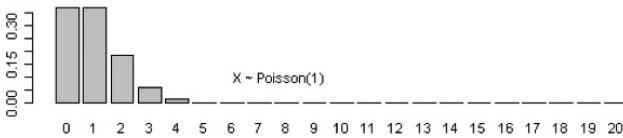
Poisson distribution: probability distribution of the Poisson random variable  $X$ , p.f. given by:

$$f_X(x) = \Pr(X=x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad \text{for } x=0, 1, 2, 3, \dots$$

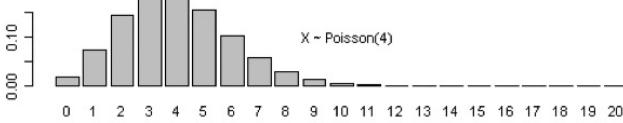
where  $\lambda$ : average number of successes occurring in the given time interval or specified region

and  $e \approx 2.718281818\dots$

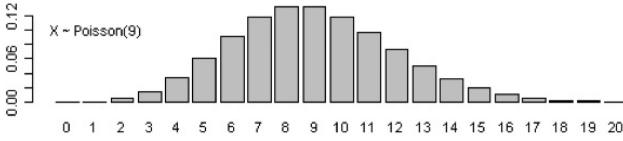
The probability distribution of a Poisson random variable



The probability distribution of a Poisson random variable



The probability distribution of a Poisson random variable



X with parameter  $\lambda$   
↗

- Theorem 4.4 : Mean and variance of a Poisson random variable :

$$E(X) = \lambda$$

$$V(X) = \lambda$$

## 4.5 - Poisson Approximation to the Binomial Distribution

- **Theorem 4.5:** Let  $X$  be a binomial random variable with parameters  $n$  and  $p$ ,

i.e.  $\Pr(X=x) = f_X(x) = \binom{n}{x} p^x q^{n-x}$ , where  $q = 1-p$ .

Suppose  $n \rightarrow \infty$  and  $p \rightarrow 0$  in such a way that  $\lambda = np$  remains a constant as  $n \rightarrow \infty$ .

Then  $X$  will approximately have Poisson distribution with parameter  $np$ , i.e.  $\lim_{n \rightarrow \infty} \Pr(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}$

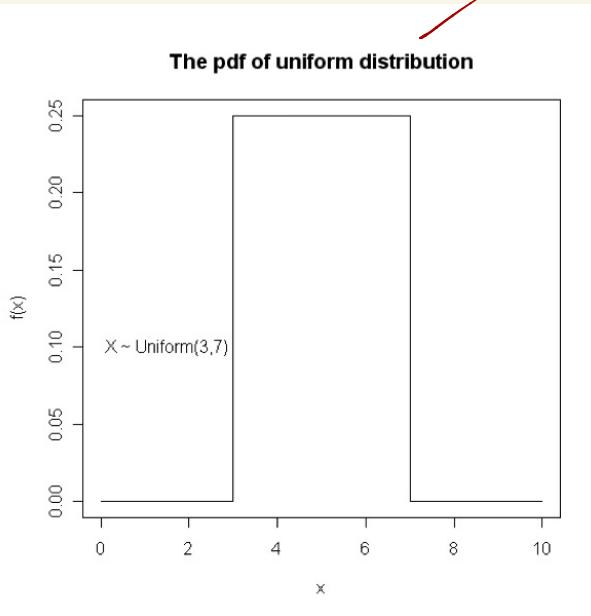
- If  $p$  is close to 1, we can still use Poisson distribution to approximate binomial probabilities by interchanging what we have defined to be a success and a failure so that changing  $p$  to a value close to zero

## 4.6 - Continuous Uniform Distribution

- Definition 4.5 : A random variable has a uniform distribution over the interval  $[a, b]$ ,  $-\infty < a < b < \infty$ , denoted by  $U(a, b)$ , if its p.d.f. is given by:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b; \\ 0 & \text{otherwise.} \end{cases}$$

a.k.a. rectangular distribution



- Theorem 4.6 : Mean and variance of a continuous uniform random variable :

If  $X$  is uniformly distributed over  $[a, b]$ , then :

$$E(X) = \frac{a+b}{2}$$

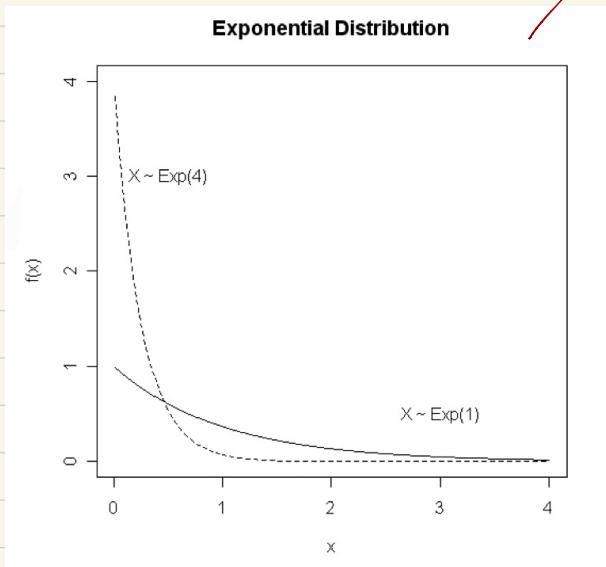
$$V(X) = \frac{1}{12}(b-a)^2$$

## 4.7 - Exponential Distribution

- Definition 4.6 : A continuous random variable  $X$  assuming all non-negative values has an **exponential distribution** with parameter  $\alpha > 0$  if its p.f. is given by:

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Note:  $\int_{-\infty}^{\infty} f(x) dx = 1$



- Theorem 4.7 : Mean and variance of an exponential random variable :

If  $X$  has an exponential distribution with parameter  $\alpha > 0$ , then:

$$E(X) = \frac{1}{\alpha}$$

$$V(X) = \frac{1}{\alpha^2}$$

- p.d.f. can be written in the form:  $f_X(x) = \begin{cases} \frac{1}{\mu} e^{-\frac{x}{\mu}} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$

Then  $E(X) = \mu$  and  $V(X) = \mu^2$ .

- Theorem 4.8 : Suppose  $X$  has an exponential distribution with parameter  $\alpha > 0$ .

"No memory"  
property of  
exponential  
distribution

Then for any two positive numbers  $s$  and  $t$ , we have :

$$\Pr(X > s+t \mid X > s) = \Pr(X > t)$$

- Application of exponential distribution : frequently used as a model for the distribution of times between the occurrence of successive events e.g. customers arriving at a service facility or calls coming into a switchboard

## 4.8 - Normal Distribution

$-\infty < x < \infty$

- Definition 4.7 : The random variable  $X$  assuming all **real values** has a

normal (Gaussian) distribution if its p.d.f. is given by:  

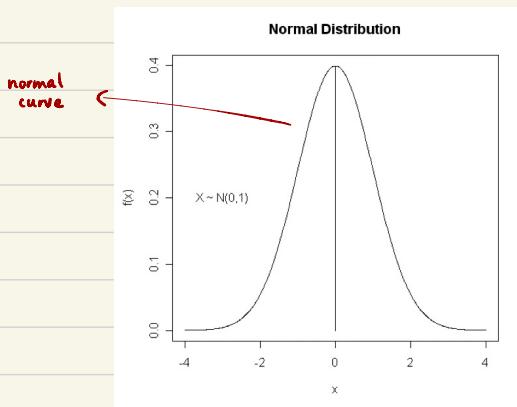
$$X \sim N(\mu, \sigma^2) \quad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty,$$

where  $-\infty < \mu < \infty$  and  $\sigma^2 > 0$ .

parameters of  
the normal  
distribution

- Properties of the normal distribution :

① Graph is bell-shaped, symmetrical about vertical line  $x = \mu$



② Maximum point occurs at  $x = \mu$ , value:  $\frac{1}{\sqrt{2\pi}\sigma}$

③ Normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean

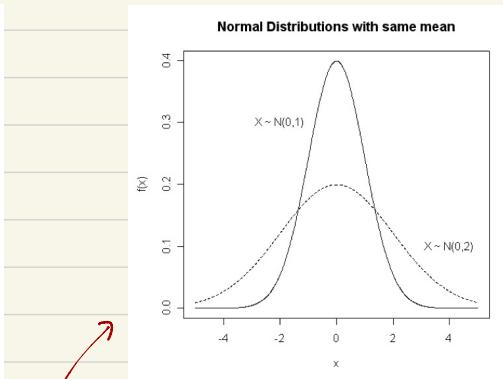
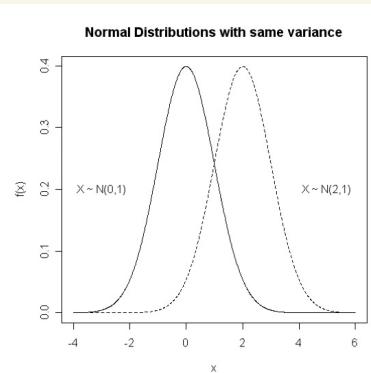
④ Total area under the curve and above the horizontal axis is equal to 1  

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1$$

⑤  $E(X) = \mu$  and  $V(X) = \sigma^2$

⑥ Two normal curves are identical in shape if they have the same  $\sigma^2$ , but they are centered at different positions when their means are different

(b)



- ⑦ As  $\sigma$  increases, the curve flattens; and as  $\sigma$  decreases, the curve sharpens
- ⑧ If  $X$  has distribution  $N(\mu, \sigma^2)$ , and if  $Z = \frac{(X-\mu)}{\sigma}$ , then  
     $Z$  has the  $N(0, 1)$  distribution. i.e.  $E(Z) = 0$  and  $V(Z) = 1$
- ∴  $Z$  has a standardized normal distribution and p.d.f. of  $Z$  is:
- $$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$
- Let  $z_1 = \frac{x_1-\mu}{\sigma}$  and  $z_2 = \frac{x_2-\mu}{\sigma}$ . Then
- $$\Pr(x_1 < X < x_2) = \Pr(z_1 < Z < z_2)$$
- Standardized normal distribution is tabulated  $\Rightarrow$  useful!

// TODO  
Statistical tables?

## 4.9 - Normal Approximation to the Binomial Distribution