Binominal Distribution & Poisson Distribution

Huining Kang

HuKang@salud.unm.edu

July 29, August 3, 2020

Overview

- Contents
  - Review probability distribution concepts
  - Binomial distribution
    - Bernoulli trial
    - Bernoulli process
    - Binomial distribution
    - Probability calculation based on binomial distribution
  - Poisson distribution
    - Definition
    - The Poisson Process
    - When to use Poisson Model
    - Examples

Overview

- Objective
  - Consolidate the understanding of the concepts related to probability distribution
  - Understand the Binomial distribution and how to use it to calculate probabilities in real-world problems.
  - Understand the Poisson distribution and how to use it to calculate probabilities in real-world problems.

Probability Distribution (Review)

- Example 4.2.1

# of Programs	1	2	3	4	5	6	7	8	Total
Frequency	62	47	39	39	58	37	4	11	297

  - Experiment: Pick one family at random, look at the # of programs (X) the family used in the past 12 months
  - Probability distribution: Relationship between the outcomes and the probability of their occurrence. ($outcome \rightarrow probability$)

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

  - $p(x)$ is called probability density function (*pdf*) or probability mass function (*pmf*)
  - $F(x)$ is called cumulative distribution function (*cdf*)

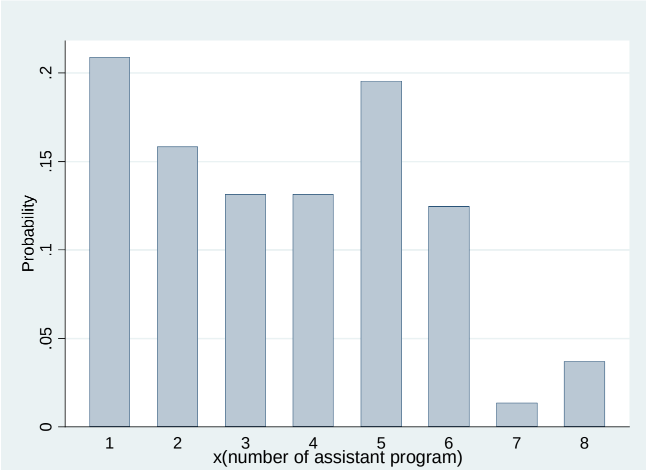
Note. Sometimes we use $f(x)$ rather then $p(x)$ to denote the *pdf*

Probability Distribution (Review)

- Regarding $p(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370

- Graphical presentation



- formula

$$p(x) = \begin{cases} .2088 & if\ x = 1, \\ .1582 & if\ x = 2, \\ .1313 & if\ x = 3, \\ .1313 & if\ x = 4, \\ .1953 & if\ x = 5, \\ .1246 & if\ x = 6, \\ .0135 & if\ x = 7, \\ .0370 & if\ x = 8. \end{cases}$$

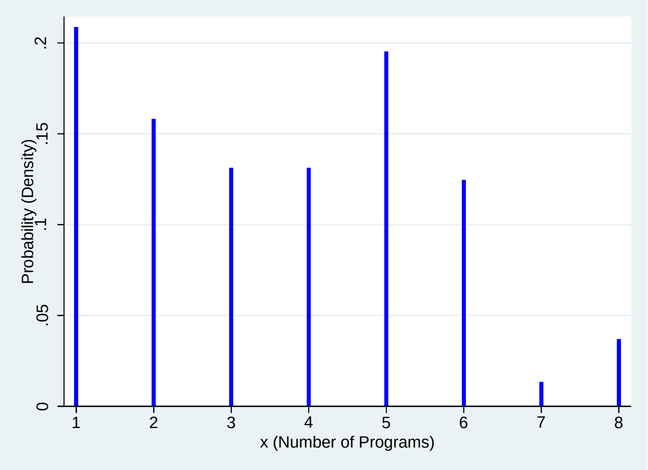
- Properties
  - $0 \leq p(x) \leq 1,\ 0 \leq P(X = x) \leq 1,$
  - $\sum p(x) = 1, \sum P(X = x) = 1.$

Probability Distribution (Review)

- Regarding $p(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370

- Graphical presentation



- formula

$$p(x) = \begin{cases} .2088 & if\ x = 1, \\ .1582 & if\ x = 2, \\ .1313 & if\ x = 3, \\ .1313 & if\ x = 4, \\ .1953 & if\ x = 5, \\ .1246 & if\ x = 6, \\ .0135 & if\ x = 7, \\ .0370 & if\ x = 8. \end{cases}$$

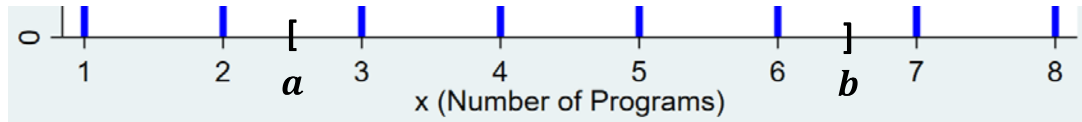
- Properties
  - $0 \leq p(x) \leq 1,\ 0 \leq P(X = x) \leq 1,$
  - $\sum p(x) = 1, \sum P(X = x) = 1.$

Probability Distribution (Review)

- Regarding $F(x)$
  - Interval events

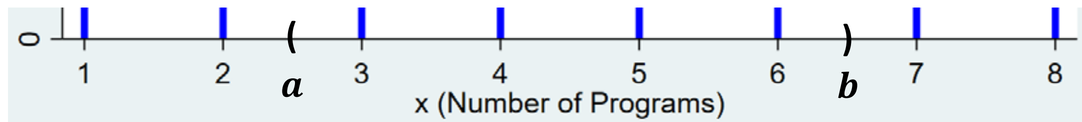
$a \leq X \leq b$

X belongs to $[a,\ b]$ (inclusively)



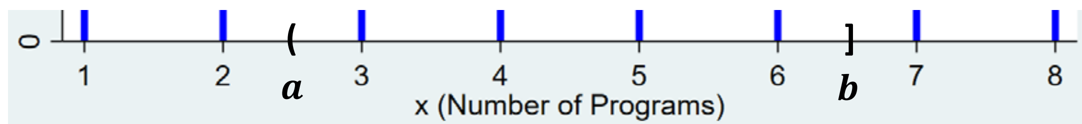
$a < X < b$

X belongs to $(a,\ b)$ (exclusively)



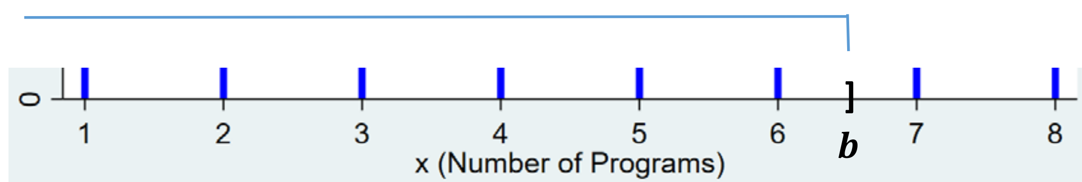
$a < X \leq b$

X belongs to $(a,\ b]$ (half-open)



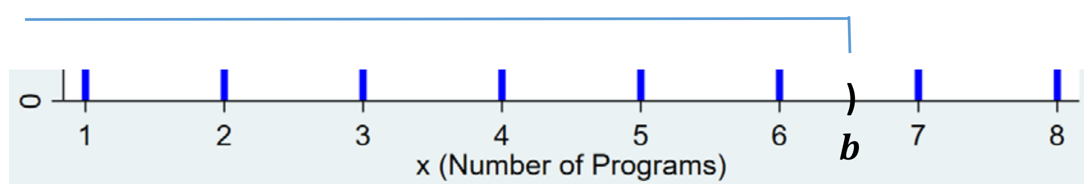
$X \leq b$

X belongs to $(-\infty,\ b]$



$X < b$

X belongs to $(-\infty,\ b)$



Probability Distribution (Review)

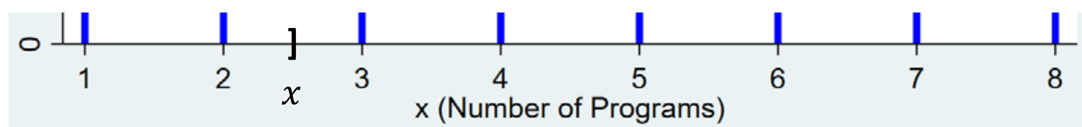
- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

- Understanding of $F(x) = P(X \leq x)$

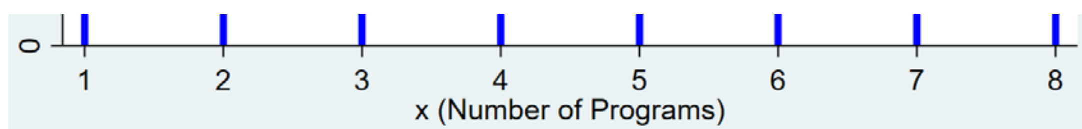
$X \leq x$

X belongs to $(-\infty,\ x]$



$F(x) = P(X \leq x) =$ Probability that X belongs to $(-\infty,\ x]$

- Example



$F(x) = P(X \leq x) =$

Probability Distribution (Review)

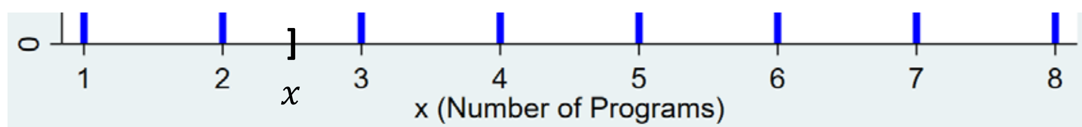
- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

  - Understanding of $F(x) = P(X \leq x)$

$X \leq x$

X belongs to $(-\infty,\ x]$



$F(x) = P(X \leq x) = 0$, if $x < 1$.

$F(1) = P(X \leq 1) = p(1) = .2088$.

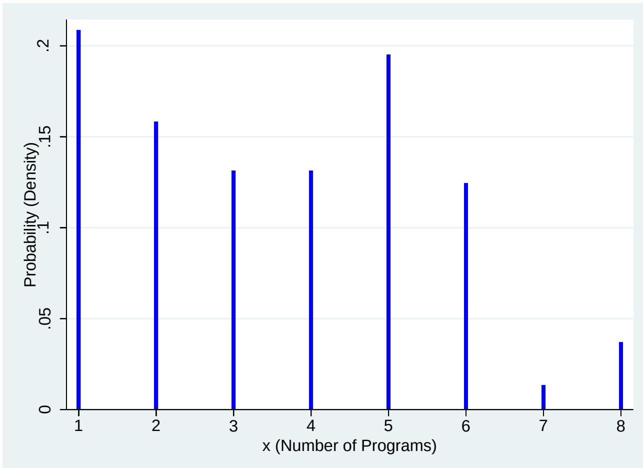
$F(x) = P(X \leq x) = p(1) = .2088$, if $1 \leq x < 2$.

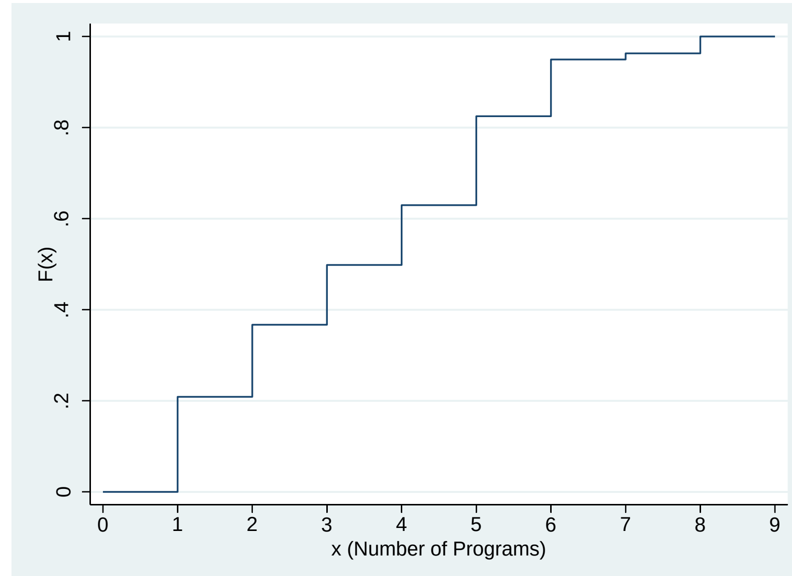
$F(2) = P(X \leq 2) = p(1) + p(2) =$ .2088+.1582 = .3670

$F(x) = P(X \leq x) = p(1) + p(2) =$ .2088+.1582 = .3670, if $2 \leq x < 3$.

$\vdots$ $\vdots$

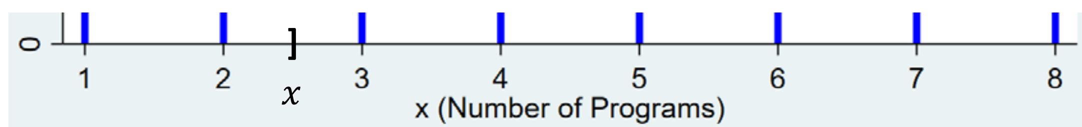
Probability Distribution (Review)





$X \leq x$

X belongs to $(-\infty,\ x]$



$F(x) = P(X \leq x) = 0$, if $x < 1$.

$F(1) = P(X \leq 1) = p(1) = .2088.$

$F(x) = P(X \leq x) = p(1) = .2088$, if $1 \leq x < 2$.

$F(2) = P(X \leq 2) = p(1) + p(2) =$ .2088+.1582 = .3670

$F(x) = P(X \leq x) = p(1) + p(2) =$ .2088+.1582 = .3670, if $2 \leq x < 3$.

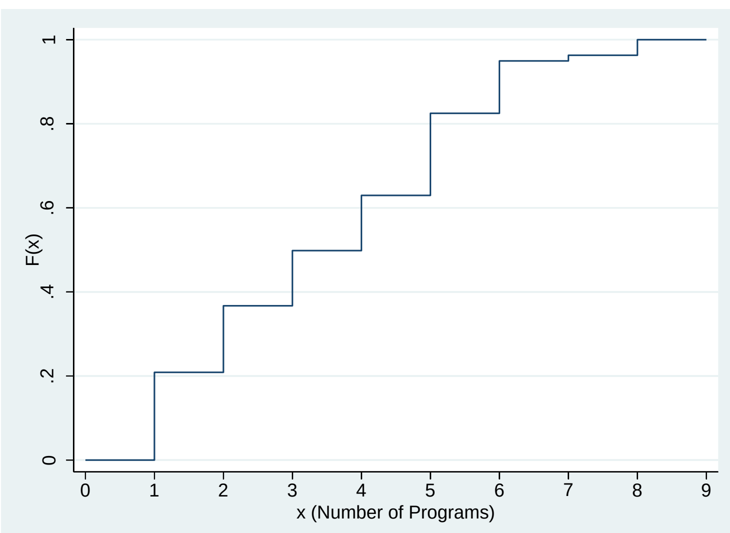
$\vdots \qquad \vdots$

Probability Distribution (Review)

- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

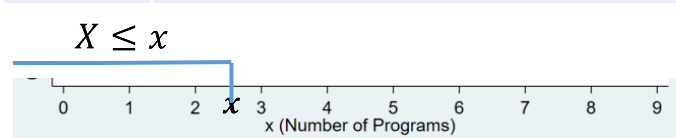
- Graphical presentation



- Understanding

x	Event "$X \leq x$"	$F(x) = P(X \leq x)$
0	$X \leq 0$ is a null event	0
0.99	$X \leq 0.99$ is a null event	0
< 1	"$X < x$" is null	0
1	"$X \leq 1$" $=$ ("$X < 1$" ∪ "$X = 1$")	.2088
1.5	"$X < 1$" ∪ "$X = 1$"∪ "$1 < X \leq 1.5$"	.2088
< 2	"$X < 1$" ∪ "$X = 1$"∪ "$1 < X \leq x$"	.2088
2	"$X < 2$" ∪ "$X = 2$"	.2088+.1582
…	…	…

- Interpretation of the event

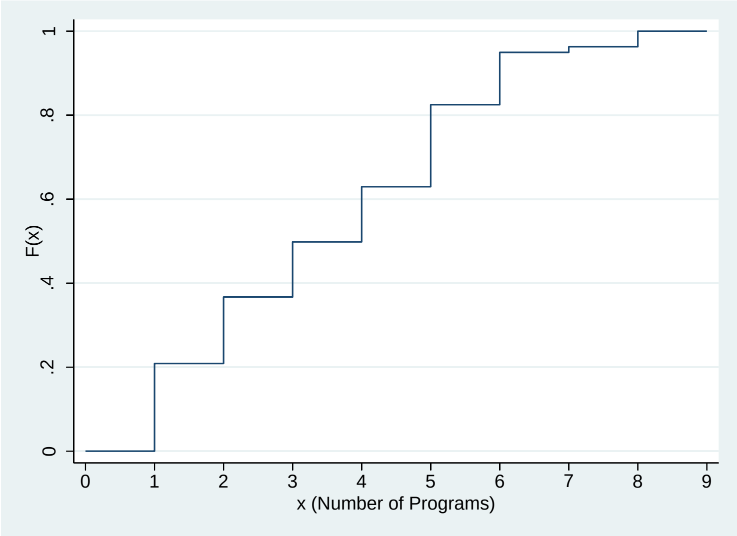


Probability Distribution (Review)

- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

- Graphical presentation



- Formula

$$F(x) = \begin{cases} 0 & if\ x < 1, \\ .2088 & if\ 1 \leq x < 2, \\ .3670 & if\ 2 \leq x < 3, \\ & \vdots \\ & \vdots \\ & \vdots \\ .9630 & if\ 7 \leq x < 8, \\ 1 & if\ x \geq 8. \end{cases}$$

Probability Distribution (Review)

- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

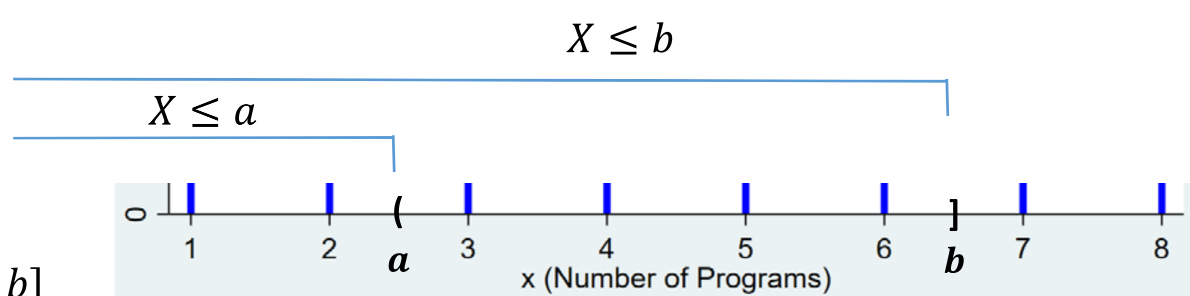
- Some useful formulae

$$P(a < X \leq b) = F(b) - F(a)$$

$$\text{“}X \leq b\text{”} = \text{“}X \leq a\text{”} \cup \text{“}a < X \leq b\text{”}$$

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

$$F(b) = F(a) + P(a < X \leq b)$$



$a < X \leq b$

X belongs to $(a,\ b]$

Probability Distribution (Review)

- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \leq x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

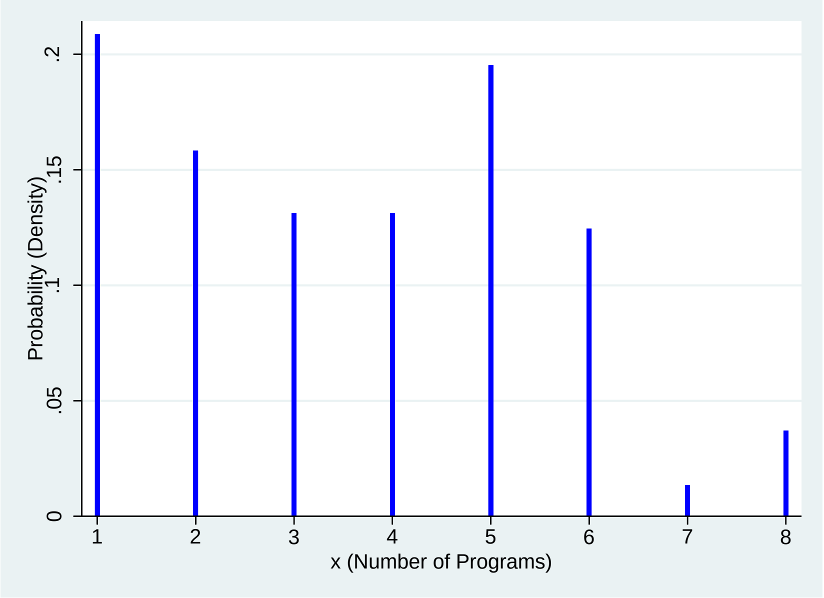
- Some useful formulae

$$P(a < X \leq b) = F(b) - F(a)$$

- Examples

$$P(3 < X \leq 6) = F(6) - F(3)$$
$$= .9495 - .4983 = .4512$$

$$P(3 \leq X \leq 6) = P(2 < X \leq 6)$$
$$= F(6) - F(2) = .9495 - .3670$$
$$= .5825$$



Probability Distribution (Review)

- Regarding $F(x)$

Posible outcome (X)	x							
	1	2	3	4	5	6	7	8
Probability $p(x) = P(X = x)$	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$
	.2088	.1582	.1313	.1313	.1953	.1246	.0135	.0370
Cumulative probability $F(x) = P(X \le x)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$
	.2088	.3670	.4983	.6296	.8249	.9495	.9630	1

- Some useful formulae

$$P(X > x) = 1 - F(x)$$

"$X \le x$" is the complement of "$X > x$"

$$P(X > x) = 1 - P(X \le x)$$

$$p(x) = P(X = x) = F(x) - F(x - 1)$$

Bernoulli trial

- Definition
  - A random experiment that has one of two mutually exclusive outcomes called Success (S) and Failure (F) with $p(S) = p$ and $p(F) = q$, where $p + q = 1$.
- Examples
  - Tossing a fair coin: $S =$ "head up", $F =$ "tail up", $p = 0.5$, $q = 0.5$.
  - Tossing a die  : $S =$ "six dots", $F =$ "less than six dots", $p = 1/6$, q =5/6.
  - Five year survival rate of patients with certain cancer is 0.85, draw one patient from the population with the cancer. $S =$ "alive for five and more years", $F =$ "died within five years", $p = .85$, $q = .15$.
- Random variable

$$X = \begin{cases} 0, & if\ F\ occurs, \\ 1, & if\ S\ occurs, \end{cases}$$

X	0	1
P	q	p

$$p(x) = \begin{cases} q, & if\ x = 0 \\ p, & if\ x = 1 \end{cases} \quad F(x) = \begin{cases} 0, & if\ x < 0, \\ q, & if\ 0 \le x < 1, \\ 1, & if\ x \ge 1. \end{cases}$$

Bernoulli Process and Binomial Distribution

- Definition
  - A sequence of (*n*) independent Bernoulli trails with p remain constant from trial to trail.
- Examples
  - Tossing a fair coin for n times: $S =$ "head up", $F =$ "tail up".
  - A sample of *n* patients from a population of certain disease, each is either dead or alive. *S* = "dead", *F* = "Alive". $p = P(S)$ is the mortality.
- Binomial random variable
  - Let X be the number of successes in the Bernoulli process (the *n* independent Bernoulli trails). Then the possible outcome is $X = 0, 1, 2, \cdots, n$
  - The probability distribution of X is called Binomial distribution

Binomial Distribution

- Basic setting
  - A sequence of (n) independent Bernoulli trails.
  - Let S_i and F_i to denote the success and failure events in the ith trial, $(i = 1, 2, \cdots, n)$.
  - $P(S_i) = p, P(F_i) = q$ are the same for all the trials.
  - Let X be the number of successes in the n. Then $X = 0, 1, 2, \cdots, n$.
- Question of interest
  - What is the probability

$$f(x) = P(X = x), x = 0, 1, \cdots n \,?$$

- When $n = 1$
  -

X	0	1
P	q	p

$$f(x) = \begin{cases} q, & if\ x = 0, \\ p, & if\ x = 1. \end{cases}$$

Binomial Distribution

- Basic setting
  - A sequence of (n) independent Bernoulli trails.
  - Let S_i and F_i to denote the success and failure events in ith trial, $(i = 1, 2, \cdots, n)$.
  - $P(S_i) = p, P(F_i) = q$ are the same for all the trials.
  - Let X be the number of successes in the n. Then $X = 0, 1, 2, \cdots, n$.
- When $n = 2$
  - Four possible outcomes:

$$\underbrace{F_1 \cap F_2}_{X=0},\ \underbrace{S_1 \cap F_2,\ F_1 \cap S_2}_{X=1},\ \underbrace{S_1 \cap S_2}_{X=2}$$

$$f(0) = P(X = 0) = P(F_1 \cap F_2) = P(F_1) \cdot P(F_2) = q \cdot q = q^2$$

$$\begin{aligned} f(1) = P(X = 1) = P((S_1 \cap F_2) \cup (F_1 \cap S_2)) &= P(S_1 \cap F_2) + P(F_1 \cap S_2) \\ &= P(S_1)P(F_2) + P(F_1)P(S_2) = p \cdot q + q \cdot p = 2pq \end{aligned}$$

$$f(2) = P(X = 2) = P(S_1 \cap S_2) = P(S_1) \cdot P(S_2) = p \cdot p = p^2$$

- Distribution of X

X	0	1	2
P	q^2	$2pq$	p^2

Binomial Distribution

$$(a+b)^n = a^4 + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \cdots + \frac{n(n-1)\cdots(n-x+1)}{x!}a^{n-x}b^x + \cdots + nab^{n-1} + b^n$$

- Binomial Theorem

$$(a+b)^0 = 1$$

$$(a+b)^1 = a+b$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = ?a^4 + ?a^3b + ?a^2b^2 + ?ab^3 + ?b^4$$

$$(a+b)^4 = 1a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + 1b^4$$

$$\vdots$$

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
\end{array}
$$

… … … … … … … …

Pascal's triangle. Also called Yang Hui's triangle

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \cdots + \frac{n(n-1)\cdots(n-x+1)}{x!}a^{n-x}b^x + \cdots + nab^{n-1} + b^n$$

$${}_nC_x = \frac{n(n-1)\cdots(n-x+1)}{x!}$$

${}_nC_x$ is the number of combinations of n objects that can be formed by taking x of them at a time, read n choose x.

$x! = x \times (x-1) \times (x-2) \times \ldots \times 2 \times 1$, is called x factorial

Binomial Distribution

- Basic setting
  - A sequence of (n) independent Bernoulli trails.
  - Let S_i and F_i to denote the success and failure events in ith trial, $(i = 1, 2, \cdots, n)$.
  - $P(S_i) = p, P(F_i) = q$ are the same for all the trials.
  - Let X be the number of successes in the n. Then $X = 0, 1, 2, \cdots, n$.
- Probability Distribution of X

X	0	1	$\cdots$	x	$\cdots$	$n-1$	n
P	q^n	npq^{n-1}	$\cdots$	${}_nC_x p^x q^{n-x}$	$\cdots$	$np^{n-1}q$	p^n

$$P(X = x) = f(x) = {}_nC_x p^x q^{n-x}, \quad x = 0, 1, \cdots, n$$

$$q^n + npq^{n-1} + \cdots + {}_nC_x p^x q^{n-x} + \cdots + np^{n-1}q + p^n = (q + p)^n = 1$$

Binomial Distribution

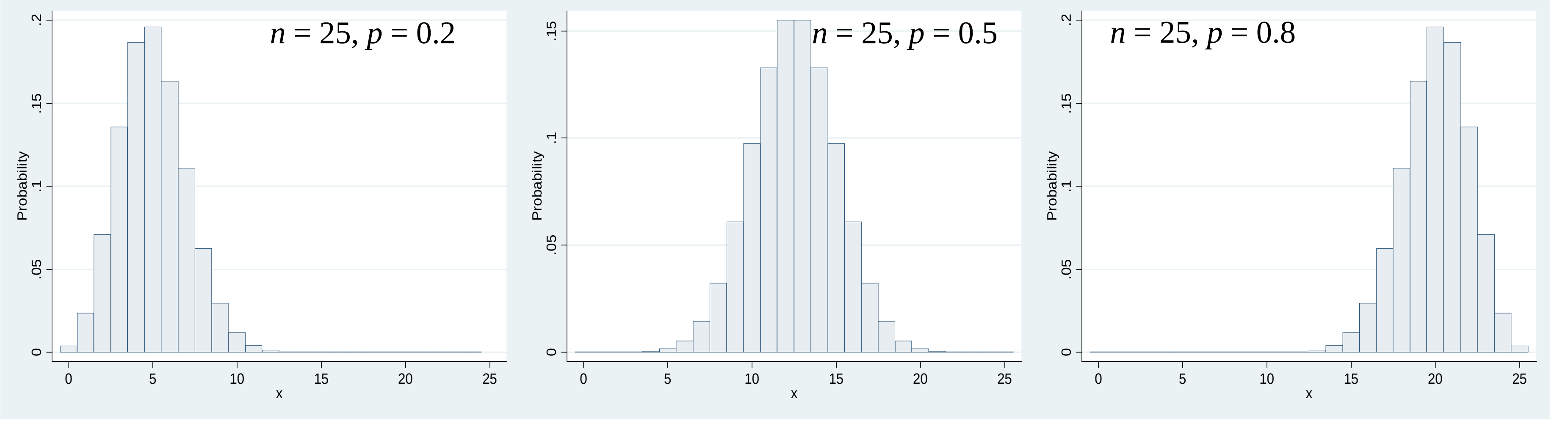
- Calculation of Binomial Distribution
  - Calculation with cdf $P(x_1 < X \le x_2) = F(x_2) - F(x_1)$
    especially for pdf $P(X = x) = F(x) - F(x-1)$
  - Binomial Table (Table B, page A-3 through A-31)
  - Stata functions
    - binomialp(n, x, p) for calculating pdf $f(x) = P(X = x)$
    - binomial(n, x, p) for calculating $F(x) = P(X \le x)$
    - comb(n, x) for calculating ${}_nC_x$
- Issue about using the table when $p > .5$

Binomial distribution

- Example 4.3.3 (Page 103)
  - 10% of a certain population is color blind
  - Draw a random sample of 25 people from the population
- Probability Calculation
  - S = "color blind", F = "not color blind", $p = P(S) = 0.1$, $q = 0.9$, $n = 25$
  - (a) Five or fewer will be color blind : $P(X \leq 5) = F(5)$
    - disp binomial(25, 5, 0.1)
    - .9666
  - (b) Six or more will be color blind: $P(X \geq 6) = 1 - P(X \leq 5) = .0334$
  - (c) Between six and nine inclusive will be color blind:
    - $P(6 \leq X \leq 9) = P(X \leq 9) - P(X \leq 5) = ?$
    - disp binomial(25, 9, 0.1) - binomial(25, 5, 0.1)
    - 0.0333
  - (d) Two, three or four will be color blind:
    - $P(2 \leq X \leq 4) = P(X \leq 4) - P(X \leq 1) = ?$
    - disp binomial(25, 4, 0.1) - binomial(25, 1, 0.1)
    - .6308
  - (e) At least one is color blind:
    - $P(X \geq 1) = 1 - P(X = 0) = 1 - q^{25} = 1 - 0.9^{25} = 0.9282$
    - disp 1 – 0.9^25
    - disp 1 – binomial(25, 0, 0.1)
    - disp 1 – binomialp(25, 0, 0.1)

Binomial Distribution

- Often denoted by binomial(n, p) or bin(n, p)
  - n and p are two parameters
- Characteristic numbers
  - Mean (mathematical expectation): $\mu = np$
  - Variance: $\sigma^2 = npq = np(1\text{-}p)$
- Graph of the distribution



Binomial Distribution

- R code for generating the three histograms in the previous slide

```
clear
set obs 26
gen x = _n-1
gen Prob1 = binomialp(25, x, 0.2)
gen Prob2 = binomialp(25, x, 0.5)
gen Prob3 = binomialp(25, x, 0.8)
label variable Prob1 "Probability"
label variable Prob2 "Probability"
label variable Prob3 "Probability"

twoway (bar Prob1 x, sort color(blue)
fintensity(inten10))

twoway (bar Prob2 x, sort color(blue)
fintensity(inten10))

twoway (bar Prob3 x, sort color(blue)
fintensity(inten10))
```

Poisson Distribution

- Examples
  - $X =$ The number of patients arriving in an emergency room between 10 and 11 PM
  - $X =$ The number of phone calls to a customer service center in a business day
  - $X =$ The number of meteorites greater than one meter diameter that strike Earth in a year
- What is in common
  - X is the number of occurrences of some random event in an interval of time or space (or some volume of matter).
  - Possible outcome: $X = 0, 1, 2, 3, \cdots$
- Poisson Distribution
  - Suppose the mean of X is λ, then approximately we have

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, 3, \cdots.$$

x	0	1	2	3	$\cdots$
$P(X = x)$	$e^{-\lambda}$	$e^{-\lambda} \cdot \lambda$	$e^{-\lambda} \cdot \frac{\lambda}{2}$	$e^{-\lambda} \cdot \frac{\lambda^2}{6}$	$\cdots$

Poisson Distribution

- Poisson Distribution
  - Suppose the mean of X is λ, then approximately we have

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, 3, \cdots.$$

x	0	1	2	3	$\cdots$
$P(X = x)$	$e^{-\lambda}$	$e^{-\lambda} \cdot \lambda$	$e^{-\lambda} \cdot \frac{\lambda}{2}$	$e^{-\lambda} \cdot \frac{\lambda^2}{6}$	$\cdots$

- Properties
  - $f(x) > 0, \quad \sum_x f(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = 1,$
  - $\mu = E(X) = \lambda, \quad \sigma^2 = Var(X) = \lambda.$
  - Poisson distribution is only distribution of the discrete random variables whose mean and variance are equal.

The Poisson Process

A sequence of random events to which the following statements are true

- The occurrences of the events are independent. The occurrence of an event in an interval of space or time has no effect on the probability of a second occurrence of the event in the same, or any other, interval.
- Theoretically, an infinite number of occurrences of the event must be possible in the interval.
- The probability of the single occurrence of the event in a given interval is proportional to the length of the interval.
- In any infinitesimally small portion of the interval, the probability of more than one occurrence of the event is negligible.

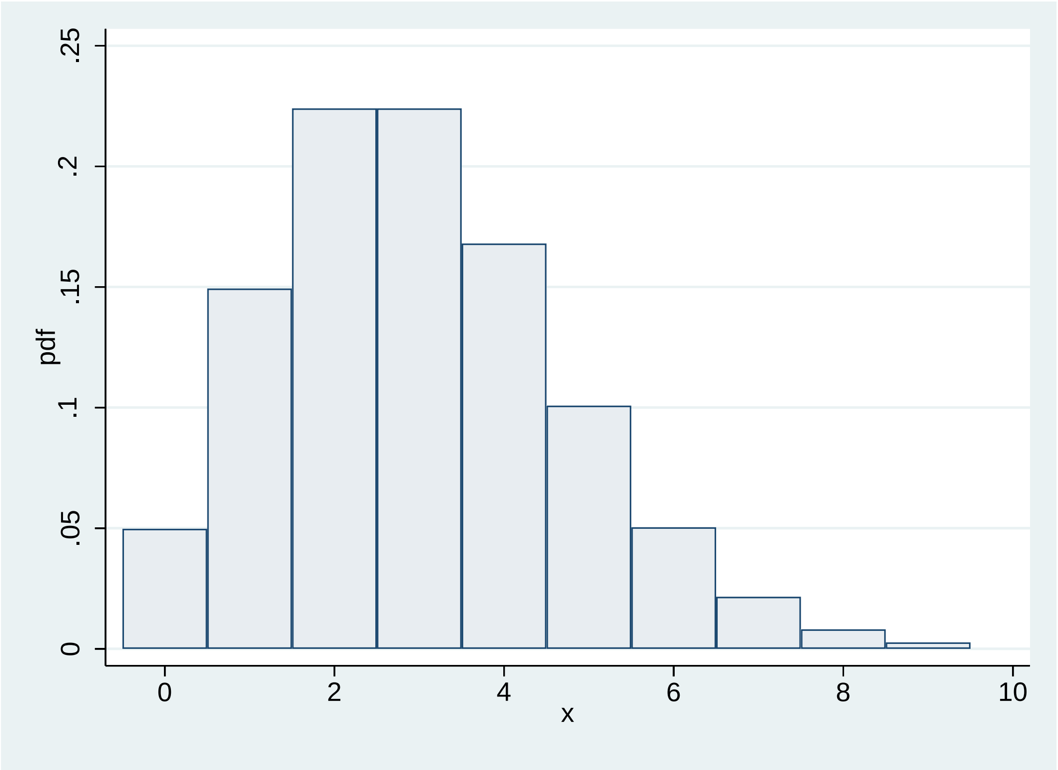
Poisson Distribution

- Calculation of Poisson Distribution
  - Calculation with cdf $P(x_1 < X \le x_2) = F(x_2) - F(x_1)$
    especially for pdf $P(X = x) = F(x) - F(x-1)$
  - Binomial Table (Table C, page A-32 through A-37)
  - Stata functions
    - Poissonp(λ, x) for calculating pdf $f(x) = P(X = x)$
    - poisson(λ, x) for calculating $F(x) = P(X \le x)$

Poisson distribution

- Example

  $\lambda = 3$



Stata code:

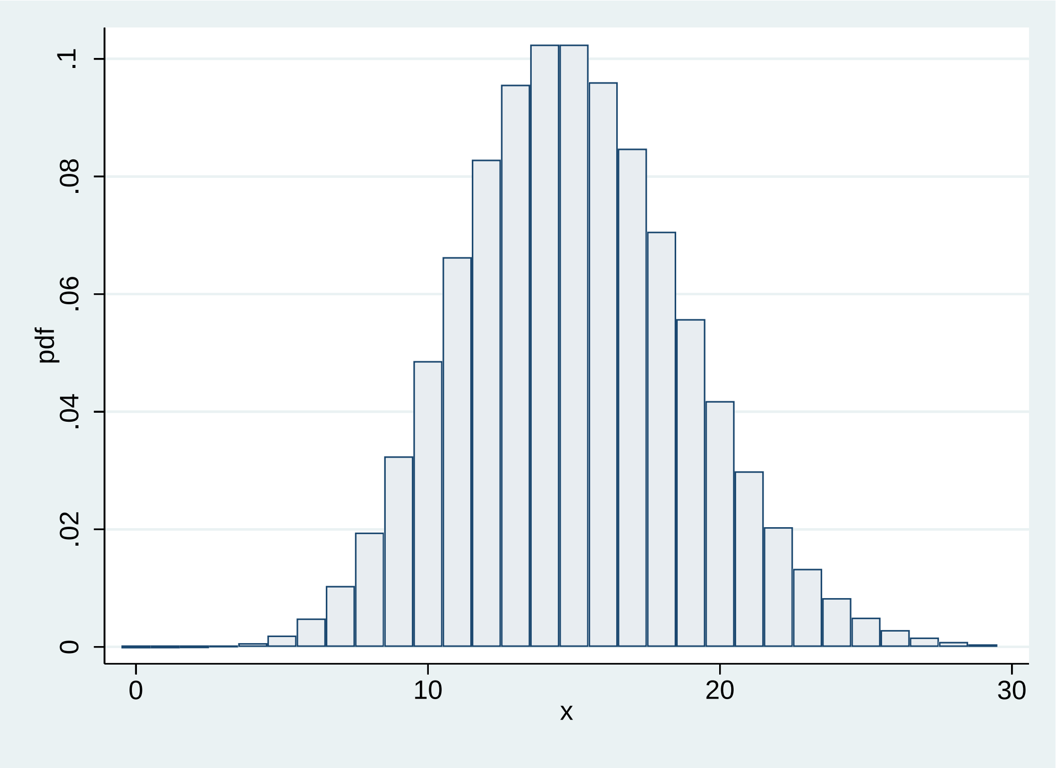
```
clear
set obs 10
gen x=_n-1
gen pdf = poissonp(3, x)
twoway (bar pdf x, fintensity(10))
```

When λ is small the distribution is skewed to the right

Poisson distribution

- Example

  $\lambda = 15$



Stata code:

```
clear
set obs 30
gen x=_n-1
gen pdf =poissonp(15, x)
twoway (bar pdf x, fintensity(10))
```

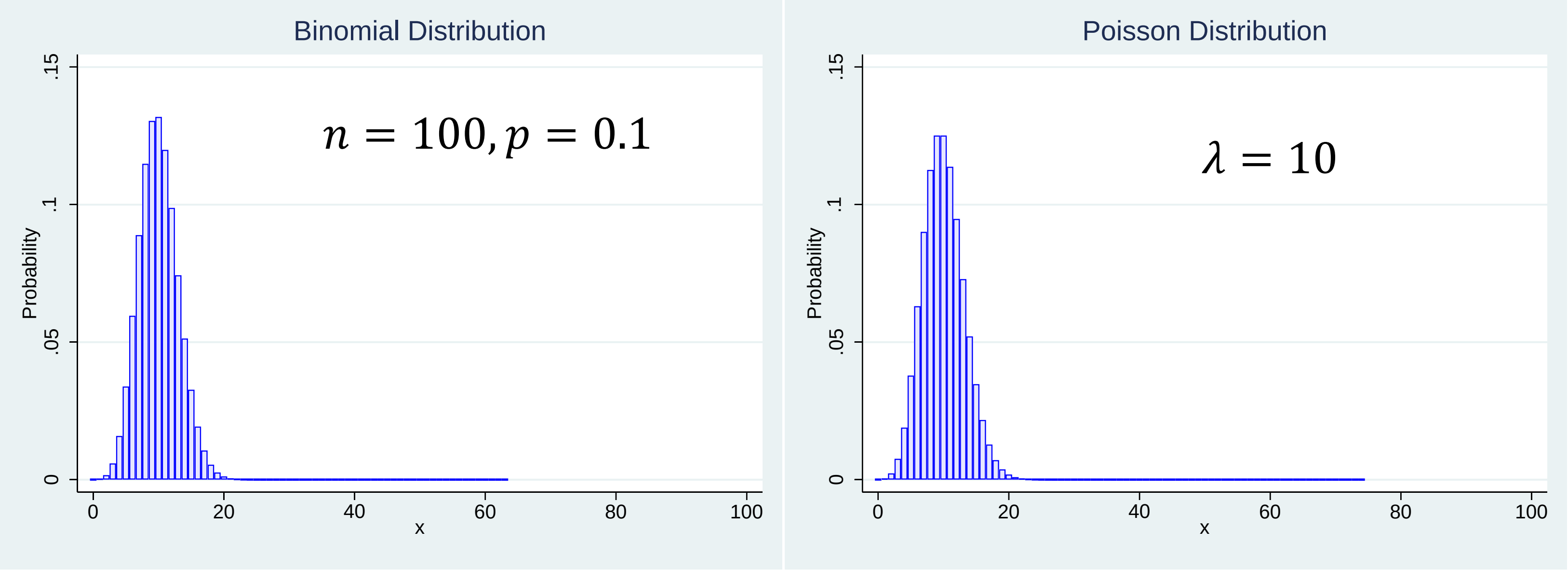
When λ is large the distribution is symmetric

Probability distributions of discrete variables – Poisson distribution

Examples (pages 110 – 112, 10^{th} Ed.)	Stata calculation
4.4.1 $\lambda = 12$, find $P(X = 3)$	. disp poissonp(12, 3) . .00176953
4.4.2 $\lambda = 12$, find $P(X \geq 3) = 1 - P(X \leq 2)$	. disp 1 – poisson(12, 2) . .99947774
4.4.3 $\lambda = 2$, find $P(X \leq 1)$	. disp poisson(2, 1) . .406
4.4.4 $\lambda = 2$, find $P(X = 3)$	. disp poissonp(2, 3) . .180
4.4.5 $\lambda = 2$, find $P(X > 5) = 1 - P(X \leq 5)$	. disp 1 – poisson(2, 5) . .01656361

Relationship between Binomial and Poisson distributions

- Suppose $X \sim Binomial(n, p)$, where n is large and p is small, and $\lambda = np$ is a moderate value. Then X is approximately Poisson distribution with mean $\lambda = np$.
- Example
  - $n = 100, p = 0.1, \lambda = np = 10$



Relationship between Binomial and Poisson distributions

- R code for comparing the Binomial and Poisson distributions

```
clear
set obs 101
gen x = _n-1
gen binP = binomialp(100, x, 0.1)
label variable binP "Probability"
gen poissonP = poissonp(10, x)
label variable poissonP "Probability"

twoway (bar binP x, sort color(blue)
fintensity(inten10)), title(Binomial Distribution)

twoway (bar poissonP x, sort color(blue)
fintensity(inten10)), title(Poisson Distribution)
```

Summary

- What we have learned
  - Review probability distribution concepts
  - Bernoulli trial
  - Bernoulli process
  - Binomial distribution
  - Probability calculation based on binomial distribution
  - Poisson distribution
  - Probability calculation based on Poisson distribution
- Learning outcomes
  - Consolidate the understanding of the concepts related to probability distribution
  - Understand the Binomial distribution and Poisson distribution and how to use them to calculate probabilities in real-world problems.
- Note
  - Understanding cdf $F(x)$ is important. Get familiar with event description and operation for a random variable.

Homework

- Due date: August 17, 2020
- Exercises 4.3.4, 4.3.5 and 4.3.6 on page 108
- Exercises 4.4.2, 4.4.4 and 4.4.5 on page 113
- Review Exercise 24 on page 131
- Review Exercises 33, 34, 35 on page 132