# STAT 3743: Probability and Statistics

G. Jay Kerns, Youngstown State University

Fall 2010

# Discrete random variables

Discrete r.v.'s have supports like

$$S_X = \{u_1, u_2, \ldots, u_k\} \text{ or } S_X = \{u_1, u_2, u_3 \ldots\}$$

Discrete r.v.'s have probability mass functions (PMFs)

$$f_X(x) = \mathbb{P}(X = x), \quad x \in S_X. \tag{1}$$

Every PMF satisfies

1. $f_X(x) > 0$ for $x \in S$,
2. $\sum_{x \in S} f_X(x) = 1$, and
3. $\mathbb{P}(X \in A) = \sum_{x \in A} f_X(x)$, for any event $A \subset S$.

# Discrete r.v.'s

**Example.**

Toss a coin four times. $X =$ number of heads

# Mean, variance, standard deviation

The mean $\mu$, a.k.a. $\mathbb{E}\,X$

$$\mu = \mathbb{E}\,X = \sum_{x \in S} x f_X(x), \tag{2}$$

The variance

$$\sigma^2 = \mathbb{E}(X - \mu)^2 = \sum_{x \in S} (x - \mu)^2 f_X(x), \tag{3}$$

The standard deviation

$$\sigma = \sqrt{\sigma^2}$$

**Example.**

Toss a coin three times. $X =$ number of heads. Find $\mu$

Interpretation:

# Cumulative distribution function (CDF)

$$F_X(t) = \mathbb{P}(X \leq t), \quad -\infty < t < \infty.$$

- $F_X$ is nondecreasing
- $F_X$ is right-continuous
- $\lim_{t \to -\infty} F_X(t) = 0$ and $\lim_{t \to \infty} F_X(t) = 1$.

Say $X$ has "distribution" $F_X$ and write $X \sim F_X$ or $X \sim f_X$

**Example.**

Toss a coin three times

# Discrete uniform distribution

$X \sim \text{disunif}(m)$ has PMF

$$f_X(x) = \frac{1}{m}, \quad x = 1, 2, \ldots, m$$

**Example.**

Roll a die

**Example.**

Select number at random from 1 to BLANK

# Find mean and variance of $X \sim \mathrm{disunif}(m)$

# Example

Find mean and variance for rolling a die

# Binomial distribution

*Bernoulli trial*: random experiment with success ($S$) and failure ($F$)

$$X = \begin{cases} 1 & \text{if the outcome is } S, \\ 0 & \text{if the outcome is } F. \end{cases}$$

Let $\mathbb{P}(S) = p$ then the PMF of $X$ is

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

$\mathbb{E}\, X$

$\mathrm{Var}(X)$

# Binomial model

- Bernoulli trials conducted $n$ times,
- the trials are independent,
- the probability of success $p$ does not change between trials.

If $X =$ number of successes then the PMF of $X$ is

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

Write $X \sim \mathrm{binom}(\mathtt{size} = n, \mathtt{prob} = p)$

# Binomial model

check $\sum f(x) = 1$:

$$\sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1$$

find the mean:

$$\mu = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= n \cdot p \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

# Example

Five child family. Let $X =$ number of boys.

**How to do it with R**

```
> dbinom(3, size = 5, prob = 0.5)
[1] 0.3125
```

# Example

Roll 15 dice. Find $\mathbb{P}(\text{from 4 to 7 sixes})$.

## How to do it with R

```
> diff(pbinom(c(4, 7), size = 15, prob = 1/6))
[1] 0.08850822
```

# Example

Seven child family, $X =$ number of boys, find CDF.

**How to do it with R**

```
> dbinom(0:7, size = 7, prob = 1/2)
[1] 0.0078125 0.0546875 0.1640625 0.2734375
[5] 0.2734375 0.1640625 0.0546875 0.0078125
```

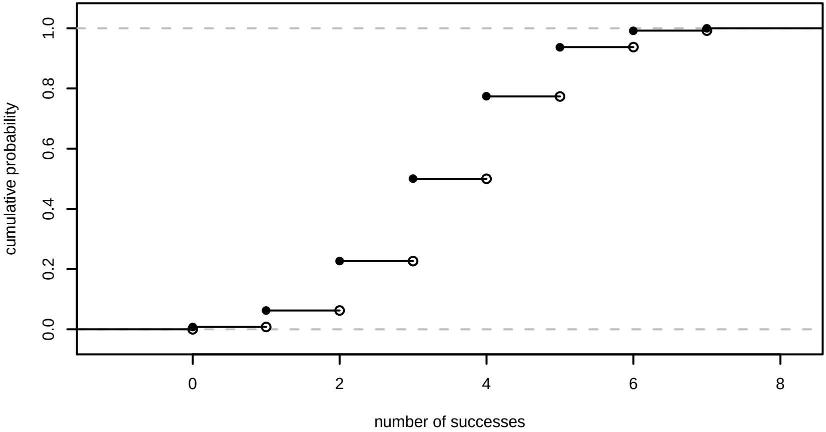


Figure: Cumulative distribution function

## Another way to do it with R

```
> library(distr)
> X <- Binom(size = 7, prob = 1/2)
> X
Distribution Object of Class: Binom
 size: 7
 prob: 0.5
> d(X)(1)    # pmf of X evaluated at x = 1
[1] 0.0546875
> p(X)(2)    # cdf of X evaluated at x = 2
[1] 0.2265625
```

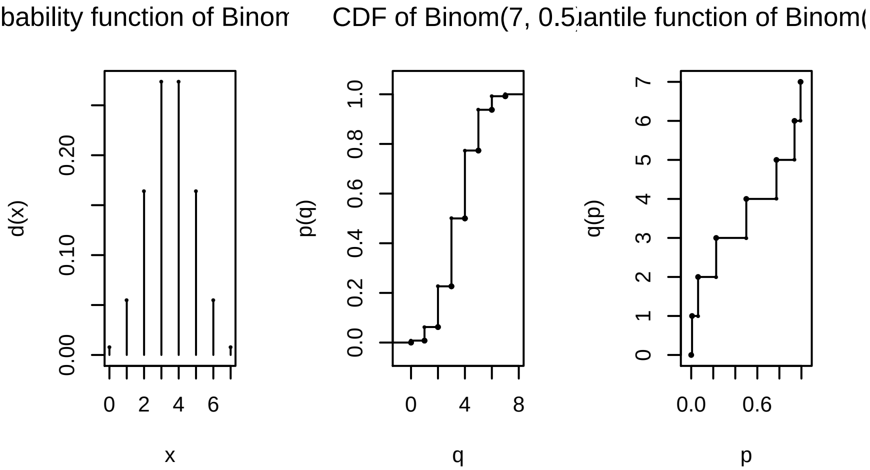


Figure: Plot of a random variable