Known Discrete Random Variables

STAT-S520

Arturo Valdivia

02-02-23

These slides complement material from ISI Chapter 4

Binomial distribution and PMF

Let $X \sim Binomial(n, p)$

- The PMF of X is

$P(X=x)$

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

for $x = 1, 2, \ldots, n$, and $f(x) = 0$ for other values of x.

PMF of binomial disribution using R

In R we use the function `dbinom()`. For the birthday party example $X \sim Binomial(\overset{n}{50}, \overset{p}{0.8})$. The probability that exactly 30 friends attend the party:

$$\underline{f(30)} = \overset{P(X=x)}{P(X = 30)} = \underbrace{\binom{50}{30} 0.8^{30}(1-0.8)^{50-30}}.$$

can be calculated in R as follows:

```
dbinom(x = 30, size = 50, prob = 0.8)
```

PMF x n p

```
## [1] 0.0006117722
```

The argument `x` in the R function is the same x that appears in our $f(x)$ math notation. The arguments `size` and `prob` correspond to the n and p parameters in the binomial distribution, respectively.

PMF of binomial disribution using R

As long as you enter the arguments in the proper order, you do not need to include the argument names, simply the values. Observe how both lines below give the same result:

```
dbinom(x = 30, size = 50, prob = 0.8)
```

```
## [1] 0.0006117722
```

```
dbinom(30, 50, 0.8)
```

```
## [1] 0.0006117722
```

The CDF of the binomial distribution

The CDF is given by

$$\underbrace{F(y) = \sum_{x:x\in L(y)} f(x)}$$

where $L(y)$ is the set of values in $X(S)$ that are less than or equal to y. If $y \in X(S)$ this is simply

$$\underbrace{F(y) = \sum_{x=0}^{y} f(x)}$$

the floor

$y = \pi$

$F(\pi) = \sum_{x=0}^{3} f(x)$

$3 = \lfloor \pi \rfloor$

CDF of binomial disribution using R

For the birthday party example $X \sim Binomial(50, 0.8)$. If we want to find the probability that 30 friends or less attend the party, then

$$\underbrace{F(30)}_{q} = P(X \leq 30) = F(30) = \sum_{x=0}^{30} f(x) = f(0) + f(1) + \cdots + f(30)$$

In R we use the function `pbinom()`:

```
pbinom(q = 30, size = 50, prob = 0.8)
```

CDF y n p

```
## [1] 0.0009324365
```

Relating PMF and CDF in R

For the b-day party example $X \sim Binomial(50, 0.8)$, there are three ways to find $F(7) = P(X \leq 7) = f(0) + f(1) + \cdots + f(7)$:

```
pbinom(7, 50, 0.8)
```

```
## [1] 1.918335e-23
```

```
dbinom(0,50, 0.8)+dbinom(1,50, 0.8)+dbinom(2,50, 0.8)+
  dbinom(3,50, 0.8)+dbinom(4,50, 0.8)+dbinom(5,50,0.8)+
  dbinom(6,50, 0.8)+dbinom(7,50, 0.8)
```

```
## [1] 1.918335e-23
```

```
sum(dbinom(0:7, 50, 0.8))
```

```
## [1] 1.918335e-23
```

Exercise 1

$[24, \infty)$

$X \geq 24$

$\underbrace{X = 0}$

$\{s \in S : X(s) = 0\}$

$\{s \in S : X(s) \geq 24\}$

For the b-day party example $X \sim \mathit{Binomial}(50, 0.8)$, find the probability that:

$1 - P(X=0) = 1 - F(0)$

- You do not celebrate alone. $1 - \text{dbinom}(0, 50, 0.5)$
- At most 35 friends attend. $F(35) = P(X \leq 35)$
- At least 24 friends attend. $P(X \geq 24) = 1 - P(X < 24) = 1 - P(X \leq 23)$
  $1 - F(23)$
- ~~No more than~~ At least 15 but less than 40 friends attend.

Let's work in R

$P(15 \leq X < 40) = P(X \leq 39) - P(X \leq 14)$

$1 - P(X \leq 0) = 1 - F(0)$

$P(X \leq 35) = F(35)$

$P(X \geq 24) = 1 - F(23)$

$P(15 \leq X < 40) = F(39) - F(14)$

Expected value, variance, and standard deviation for the binomial distribution

If $X \sim Binomial(n, p)$ then $EX = np$, $VarX = np(1-p)$, and $SD = \sqrt{np(1-p)}$.

In our b-day party example with $X \sim Binomial(50, 0.8)$ we expect for $50 \cdot 0.8 = 40$ friends to attend give or take $\sqrt{50 \cdot 0.8 \cdot 0.2} = \sqrt{8} \approx 2.83$ friends.

The geometric distribution

- X is the number of trials before a success. $X(S) = \{0, 1, \ldots\}$ then
$$X \sim geometric(p)$$
- $f(x) = p(1-p)^x$
- $EX = (1-p)/p$
- $VarX = (1-p)/p^2$
- R: `dgeom` (PMF) and `pgeom` (CDF)

PMF

CDF

The Poisson distribution

- X is a Poisson distribution with $X(S) = \{0, 1, \dots\}$ and we write $X \sim Poisson(\lambda)$
- $$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$
- $EX = \lambda$
- $VarX = \lambda$
- R: `dpois` (PMF) and `ppois` (CDF)

λ

$X \sim \text{Poisson}(3)$

$\frac{3^0 \cdot e^{-3}}{0!}$ $\frac{3^4 e^{-3}}{4!}$

$P(X \leq 4) = F(4) = f(0) + f(1) + \dots + f(4)$