Week 4: Discrete Random Variables and Probability Distributions

MATH697

Sahir Bhatnagar

September 26, 2017

McGill University

Special Discrete Distributions

1. The Binomial Distribution

Consider flipping n coins, each of which has (independent) probability p of coming up heads, and probability $1 - p$ of coming up tails, $0 < p < 1$. Let X be the total number of heads showing. The random variable X is said to have the $Bin(n, p)$ distribution with PMF given by

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

The Binomial Distribution: Expected Value and Variance

Recall that if $a, b \in \mathbb{R}$ and $n = 0, 1, 2, \ldots,$ then we have the *Binomial Formula*

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

The Binomial Distribution: Expected Value and Variance

Recall that if $a, b \in \mathbb{R}$ and $n = 0, 1, 2, \ldots,$ then we have the *Binomial Formula*

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

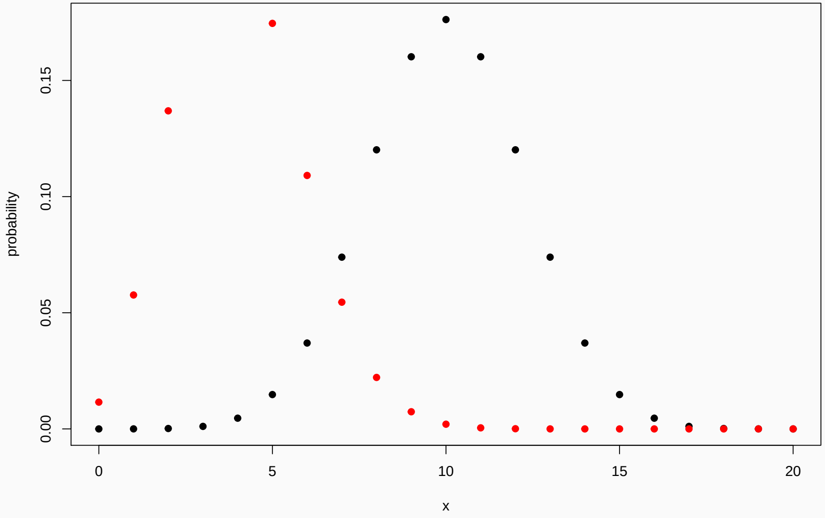
Proposition 1

$$E(X) = np$$

$$Var(X) = np(1-p)$$

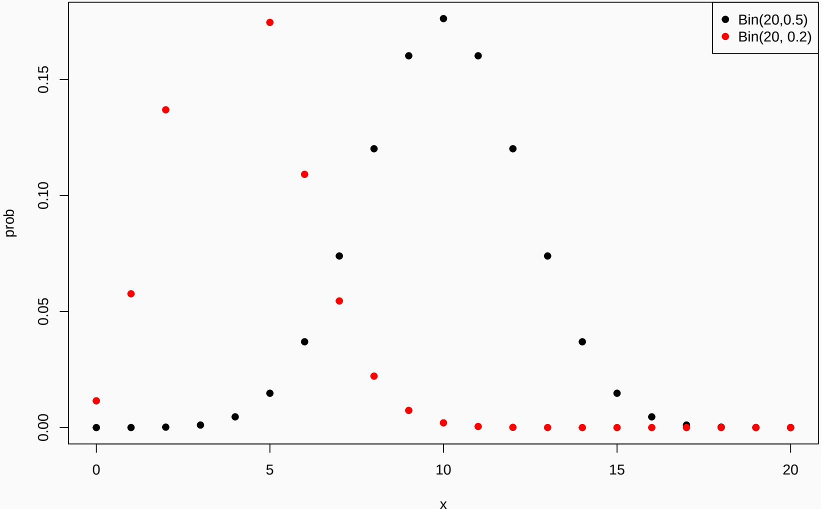
Proof: *on the board*

The Binomial Distribution: Bin(20, 0.5), Bin (20, 0.2)



The Binomial Distribution: Bin(20, 0.5), Bin (20, 0.2)

```
plot(0:20, dbinom(x = 0:20, size = 20, prob = 0.5),pch = 19,xlab="x",ylab="prob")
points(0:20, dbinom(0:20, size = 20, prob = 0.20), pch = 19, col = "red")
legend("topright", c("Bin(20,0.5)", "Bin(20, 0.2)"), col = c("black","red"), pch=19)
```



Example 2 (Exxon)

Exxon has just bought a large tract of land in northern Quebec, with the hope of finding oil. Suppose they think that the probability that a test hole will result in oil is .2. Assume that Exxon decides to drill 7 test holes. What is the probability that

1. Exactly 3 of the test holes will strike oil?
2. At most 2 of the test holes will strike oil?
3. Between 3 and 5 (including 3 and 5) of the test holes will strike oil?
4. What are the mean and standard deviation of the number of test holes which strike oil.

Binomial Distribution Exxon Example (R code)

```r
# 1
dbinom(x = 3, size = 7, prob = 0.2)
```

```
## [1] 0.114688
```

```r
# 2
pbinom(q = 2, size = 7, prob = 0.2, lower.tail = TRUE)
```

```
## [1] 0.851968
```

```r
# 3
dbinom(x = 3, size = 7, prob = 0.2) +
  dbinom(x = 4, size = 7, prob = 0.2) +
  dbinom(x = 5, size = 7, prob = 0.2)
```

```
## [1] 0.1476608
```

Binomial Distribution Exxon Example (R code)

```r
(xfx <- sapply(0:7, function(x) x * dbinom(x=x, size=7, prob=0.2))) #x*f(X=x)
```

```
## [1] 0.0000000 0.3670016 0.5505024 0.3440640 0.1146880 0.0215040 0.0021504
## [8] 0.0000896
```

```r
sum(xfx) # E(X) or E(X) = np = 7 * 0.2 = 1.4
```

```
## [1] 1.4
```

```r
(x2fx <- sapply(0:7, function(x) x^2 * dbinom(x=x, size=7, prob=0.2))) #x^2*f(X=x)
```

```
## [1] 0.0000000 0.3670016 1.1010048 1.0321920 0.4587520 0.1075200 0.0129024
## [8] 0.0006272
```

```r
sum(x2fx) - sum(xfx)^2 # V(X)
```

```
## [1] 1.12
```

```r
# or V(X) = np(1-p) = 7 * 0.2 * 0.8 = 1.12
```

2. Bernoulli Distribution

Bernoulli Distribution

The Bernoulli(p) distribution corresponds to the special case of the $Bin(n, p)$ distribution when $n = 1$, namely,

$$Bern(p) \equiv Bin(1, p)$$

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

Bernoulli Distribution

The Bernoulli(p) distribution corresponds to the special case of the $Bin(n, p)$ distribution when $n = 1$, namely,

$$Bern(p) \equiv Bin(1, p)$$

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

$X_1, X_2, ..., X_n$ are chosen independently and each has the $Bern(p)$ distribution. Then

$$Y = X_1 + \cdots + X_n \sim Bin(n, p)$$

3. Geometric Distribution

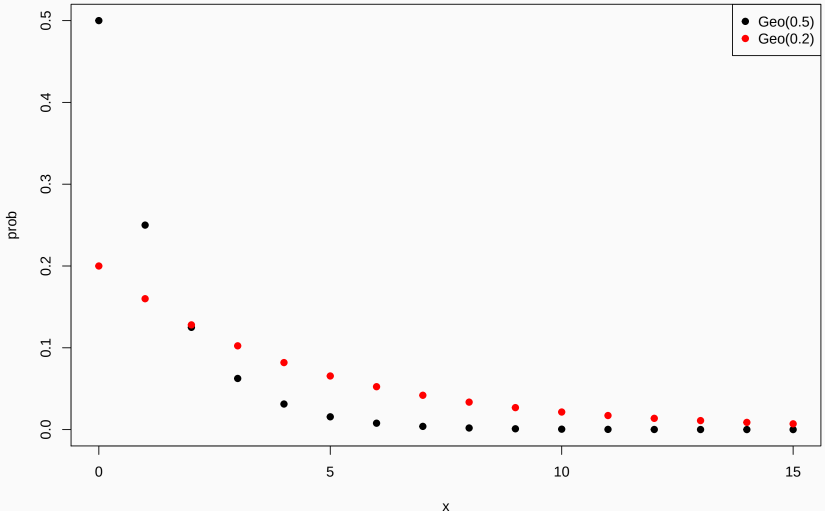
Consider repeatedly flipping a coin that has probability p of coming up heads and probability $1-p$ of coming up tails, where $0 < p < 1$. Let X be the number of tails that appear before the first head. Then for $k \geq 0$, $X = k$ if and only if the coin shows exactly k tails followed by a head. The probability of this is equal to $(1-p)^k p$

$$f_X(k) = (1-p)^k p, \quad k = 0, 1, 2, \ldots$$

We write this as $X \sim Geometric(p)$

The Geometric Distribution: Geo(0.5), Geo(0.2)

```
plot(0:15, dgeom(x = 0:15, prob = 0.5),pch = 19, xlab="x", ylab="prob")
points(0:15, dgeom(0:15, prob = 0.20), pch = 19, col = "red")
legend("topright", c("Geo(0.5)", "Geo(0.2)"), col = c("black","red"), pch=19)
```



Expected Value of Geometric Distribution

Proposition 3 (Expected Value Geo(p) distribution)

$X \sim Geo(p)$ then $E(X) = (1-p)/p$ and $V(X) = (1-p)/p^2$

Proof for E(X): *on board*

Proof for V(X): *exercise*

Alternative form of Geometric Distribution

Some books instead define the geometric distribution to be: *the first head requires k independent coin flips*. Let p be the probability of a head. Then

$$f_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, \ldots$$

4. Negative Binomial Distribution

A generalization of the geometric distribution. Consider again repeatedly flipping a coin that has probability p of coming up heads and probability $1-p$ of coming up tails. Let r be a positive integer, and let Y be the number of tails that appear before the rth head. Then for $k \geq 0$, $Y = k$ if and only if the coin shows exactly $r-1$ heads (and k tails) on the first $r-1+k$ flips, and then shows a head on the $(r+k)$-th flip. The probability of this is equal to

The Negative Binomial Distribution

A generalization of the geometric distribution. Consider again repeatedly flipping a coin that has probability p of coming up heads and probability $1 - p$ of coming up tails. Let r be a positive integer, and let Y be the number of tails that appear before the rth head. Then for $k \geq 0$, $Y = k$ if and only if the coin shows exactly $r - 1$ heads (and k tails) on the first $r - 1 + k$ flips, and then shows a head on the $(r + k)$-th flip. The probability of this is equal to

$$f_Y(k) = \binom{r-1+k}{r-1} p^{r-1}(1-p)^k p = \binom{r-1+k}{k} p^r (1-p)^k, \;\; k \in 0, 1, 2, 3,$$

The Negative Binomial Distribution

A generalization of the geometric distribution. Consider again repeatedly flipping a coin that has probability p of coming up heads and probability $1-p$ of coming up tails. Let r be a positive integer, and let Y be the number of tails that appear before the rth head. Then for $k \geq 0$, $Y = k$ if and only if the coin shows exactly $r-1$ heads (and k tails) on the first $r-1+k$ flips, and then shows a head on the $(r+k)$-th flip. The probability of this is equal to

$$f_Y(k) = \binom{r-1+k}{r-1} p^{r-1}(1-p)^k p = \binom{r-1+k}{k} p^r (1-p)^k, \ \ k \in 0, 1, 2, 3,$$

The random variable Y is said to have the Negative-Binomial(r, p) distribution.

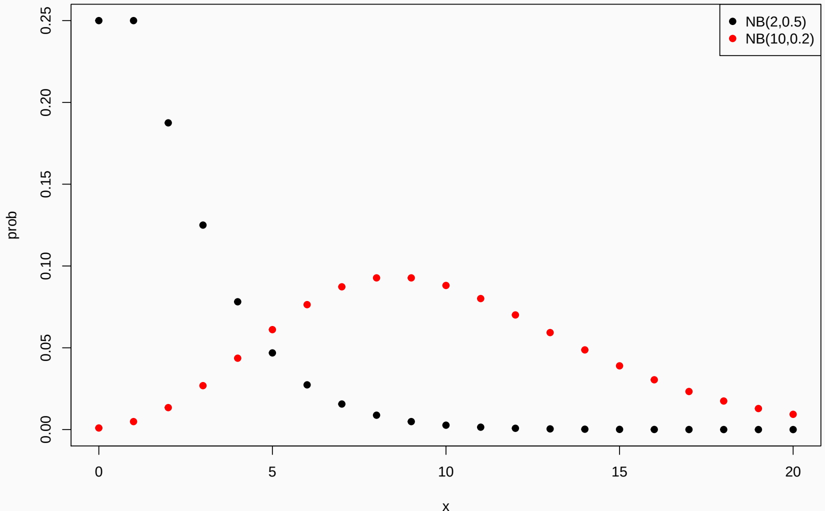
The Negative Binomial Distribution

- Let $Y \sim NegBin(r, p)$. When $r = 1$, we get then $Y \sim Geo(p)$.

- Let $Y \sim NegBin(r, p)$. When $r = 1$, we get then $Y \sim Geo(p)$.
- The Negative-Binomial(r, p) distribution applies whenever we are counting the number of failures until the rth success for independent performances of a random system where the occurrence of some event is considered a success.

The Negative Binomial Distribution: NegBin(0.5), NegBin(0.2)

```
plot(0:20, dnbinom(x = 0:20, size = 2, prob = 0.5),pch = 19, xlab="x", ylab="prob")
points(0:20, dnbinom(0:20, size = 10, prob = 0.5), pch = 19, col = "red")
legend("topright", c("NB(2,0.5)", "NB(10,0.2)"), col = c("black","red"), pch=19)
```



Example 4 (Engines)

Ten percent of the engines manufactured on an assembly line are defective. If engines are randomly selected and tested, what is the probability that

1. the first nondefective engine will be found on the second trial?
2. the third nondefective engine will be found on the fifth trial?
3. the third nondefective engine will be found on or before the fifth trial?

4. The Poisson Distribution

The Poisson Distribution

The binomial, geometric, and negative binomial distributions were all derived by starting with an experiment consisting of trials or draws and applying the laws of probability to various outcomes of the experiment. There is no simple experiment on which the Poisson distribution is based, although we will shortly describe how it can be obtained by certain limiting operations.

The Poisson Distribution

We say that a random variable Y has the *Poisson*(λ) distribution if

$$f_Y(y) = P(Y = y) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, \ldots$$

The Poisson Distribution

We say that a random variable Y has the $\mathit{Poisson}(\lambda)$ distribution if

$$f_Y(y) = P(Y = y) = \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, \ldots$$

Recall from calculus

$$e^{\lambda} = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}$$

The Poisson Distribution

We say that a random variable Y has the *Poisson*(λ) distribution if

$$f_Y(y) = P(Y = y) = \frac{\lambda^y}{y!}e^{-\lambda}, \quad y = 0, 1, 2, \ldots$$

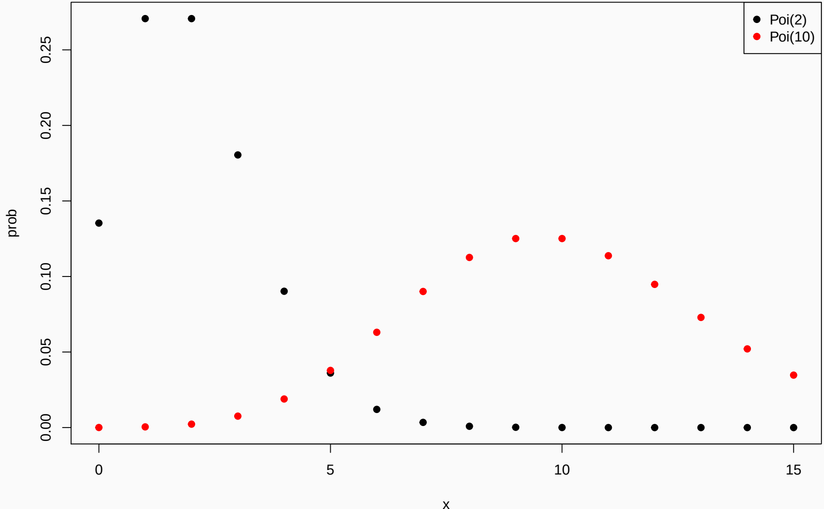
Recall from calculus

$$e^\lambda = \sum_{y=0}^{\infty} \frac{\lambda^y}{y!}$$

Exercise: show that $\sum_y f_Y(y) = 1$

The Poisson Distribution: Poi(2), Poi(10)

```
plot(0:15, dpois(x = 0:15, lambda = 2),pch = 19, xlab="x", ylab="prob")
points(0:15, dpois(0:15, lambda = 10), pch = 19, col = "red")
legend("topright", c("Poi(2)", "Poi(10)"), col = c("black","red"), pch=19)
```



The Poisson Distribution as a Limit

The rationale for using the Poisson distribution in many situations is provided by the following proposition.

Proposition 5 (Limit of a binomial is Poisson)

Suppose that $X \sim$ *Binomial*(n, p). *If we let* $p = \lambda/n$, *then as* $n \to \infty$, *Binomial*$(n, p) \to$ *Poisson*(λ). *Another way of saying this: for large n and small p, we can approximate the binomial(n,p) probability by the Poisson*$(\lambda = np)$.

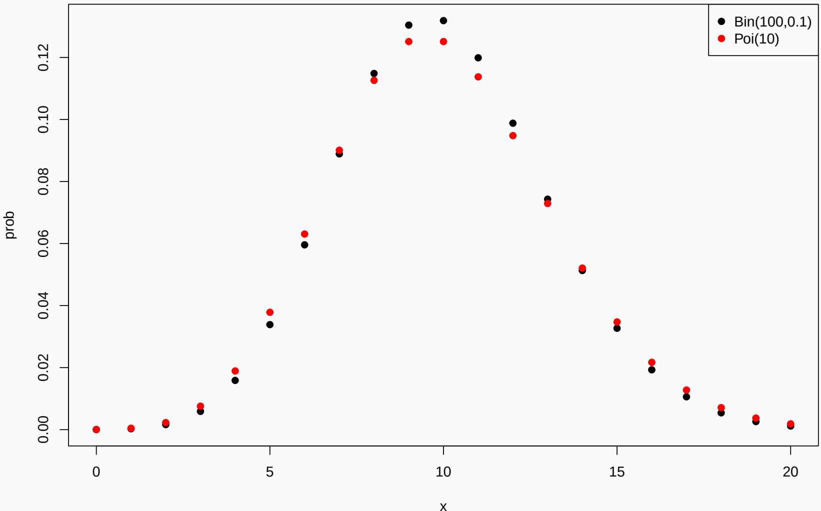
Proof: *on board*

Recall

$$\left(1 + \frac{c}{n}\right)^n \to e^c$$

Poisson approximation to the Binomial

```
plot(0:20, dbinom(x = 0:20, size = 100, p=0.1), pch = 19, xlab="x", ylab="prob")
points(0:20, dpois(0:20, lambda = 100*0.1), pch = 19, col = "red")
legend("topright", c("Bin(100,0.1)", "Poi(10)"), col = c("black","red"), pch=19)
```



Poisson Expectation and Variance

Let $X \sim Poisson(\lambda)$, then

$$E(X) = \lambda, \quad V(X) = \lambda$$

recall:

$$e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

Proof: *on board*

Summary of Discrete Distributions

Single 0-1 trial - count number of 1s	$\Longrightarrow$	BERNOULLI DISTRIBUTION
n independent 0-1 trials - count number of 1s	$\Longrightarrow$	BINOMIAL DISTRIBUTION
Sequence of independent 0-1 trials - count number of trials until first 1	$\Longrightarrow$	GEOMETRIC DISTRIBUTION
Sequence of independent 0-1 trials - count number of trials until r^{th} 1 is observed	$\Longrightarrow$	NEGATIVE BINOMIAL DISTRIBUTION
Limiting case of binomial distribution	$\Longrightarrow$	POISSON DISTRIBUTION

Moment Generating Functions

Definition 6 (Moments)

Let X be a random variable and k an integer with $k \geq 0$. Suppose that $E(|X^k|) < \infty$ (i.e. the expected value exists). Then the number $\mu'_k = E(X^k)$ is called the kth moment of X about the origin. The number $\mu_k = E[(X - \mu)^k]$ (where $\mu = \mu'_1 = E(X)$) is called the kth moment of X about its mean.

Definition 7 (MGFs)

Let X be a random variable. If there exists a $\delta > 0$, such that $E(e^{tX}) < \infty$ for all $-\delta < t < \delta$, then

$$M_X(t) \equiv E(e^{tX}), \quad -\delta < t < \delta$$

is called the MGF of X. For a discrete RV we have

Definition 7 (MGFs)

Let X be a random variable. If there exists a $\delta > 0$, such that $E(e^{tX}) < \infty$ for all $-\delta < t < \delta$, then

$$M_X(t) \equiv E(e^{tX}), \quad -\delta < t < \delta$$

is called the MGF of X. For a discrete RV we have

$$M_X(t) = \sum_{x \in R_X} e^{tx} f_X(x)$$

Example 8 (Constant)

If $X = c$, then

$$M_X(t) = E(e^{tc}) = e^{tc}$$

Example 9 (Binomial)

If $X \sim Binomial(n, p)$, then

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} = (pe^t + (1-p))^n$$

Proof: *on board*

Example 10 (Poisson)

If $X \sim Poisson(\lambda)$, then

$$M_X(t) = e^{-\lambda(1-e^t)}$$

Proof: *on board*

Why should we care about MGFs

Proposition 11

$$M_X^{(n)}(0) = E(X^n), \quad n = 0, 1, \ldots$$

In words, the n^{th} derivative of the MGF evaluated at $t = 0$ is the n^{th} moment of X

Proof.

$M_X^{(n)}(0) = E(e^0) = E(1) = 1$. Then

$$\begin{aligned} M_X'(t) &= \sum_{x \in R_X} x e^{tx} f_X(x) \\ M_X''(t) &= \sum_{x \in R_X} x^2 e^{tx} f_X(x) \\ \vdots &= \vdots \\ M_X^{(n)}(t) &= \sum_{x \in R_X} x^n e^{tx} f_X(x) \end{aligned}$$

From which $M_X'(0) = E(X)$, $M_X''(0) = E(X^2)$, and so on. □

Example 12 (Binomial)

If $X \sim Binomial(n, p)$, then

$$M_X'(t) = \frac{d}{dt}(pe^t + (1-p))^n$$
$$= n(pe^t + (1-p))^{n-1}(pe^t)$$
$$M_X''(t) = n(pe^t + (1-p))^{n-1}pe^t + n(n-1)(pe^t + (1-p))^{n-2}(pe^t)^2$$

So $E(X) = M_X'(0) = np$ and $E(X^2) = M_X''(0) = np + n(n-1)p^2$ and $V(X) = np + n(n-1)p^2 - n^2p^2 = np(1-p)$