University of Akron, Dept. of Statistics

3470:451/551 Theoretical Statistics I

Ch 3: Discrete Random Variables

Lecture Notes by Nao Mimoto

Textbook: Wackerly, Mendenhall, and Scheaffer 7e (2008)

(October 13, 2016)

Contents

2.1 Random Variable

is a function whose domain is a sample space, and whose range is a real numbers.

Discrete Random Variable is a r.v. whose range is a finite or countably infinite set.

Continuous Variable is a r.v. whose range is a interval on a real line or a disjoint union of such intervals. It also must satisfy that for any constant c, $P(X = c) = 0$.

Example

1. Throw a die: $\{1\} \to 1$
2. Throw two dice at once and add: $\{2, 5\} \to 7$

2.1.1 PMF and CDF

Probability Mass Function: (pmf)

$$p(x) = P(X = x)$$

Sum of pmf over all possible values of x must sum up to 1.

Cumulative Distribution Function: (cdf)

$$F(x) = P(X \leq x) = \sum_{t=0}^{x} p(t)$$

For any number x, $F(x)$ is the probability that the observed value of X will be at most x.

```
x <- 0:10; plot(x+.5, dbinom(x,10,.5), type="S", xlim=c(0,10), xlab="");
par(new=T)
plot(x-.5, dbinom(x,10,.5), type="h", xlim=c(0,10), xlab="pmf")

plot(x, pbinom(x,10,.5), type="s", xlim=c(0,10), xlab="CDF")
```

CDF to PDF and vice versa

If X is a discrete random variable, then

$$\begin{aligned} P(a \leq X \leq b) &= P(X \leq b) - P(X < a) \\ &= F(b) - F(a-1) \end{aligned}$$

$$P(X \geq a) = 1 - P(X < a)$$

$$\begin{aligned} P(X = a) &= P(X \leq a) - P(X \leq a-1) \\ &= F(a) - F(a-1) \end{aligned}$$

Exercises 23

Given

$$F(0) = .06 \qquad F(1) = .19 \qquad F(2) = .39$$

$$F(3) = .67 \qquad F(4) = .92 \qquad F(5) = .97 \qquad F(6) = 1$$

Calculate

a $p(2)$

$$= P(X = 2) = F(2) - F(1) = .39 - .19 = .2$$

b $P(X > 3)$

$$= 1 - P(X \leq 3) = 1 - F(3) = 1 - .39 = .61$$

c $P(2 \leq X \leq 5)$

$$= P(X \leq 5) - P(X < 2) = F(5) - F(1) = .97 - .19 = .78$$

d $P(2 < X < 5)$

$$= P(X < 5) - P(X \leq 2) = F(4) - F(2) = .92 - .39 = .53$$

2.1.2 Expected value and Variance

Expected value

- Expected Value of a random variable X, whose range is $x_1, x_2, x_3, \ldots x_n$ is defined as

$$E(X) = \mu = \sum_{i=1}^{n} x_i \cdot p(x_i)$$

- Expected Value of a function of random variable X, say $g(X)$ is defined as

$$E\big(g(X)\big) = \sum_{i=1}^{n} g(x_i) \cdot p(x_i)$$

- If a and b are constants, then

$$E(aX + b) = aE(X) + b$$

- Expectation value is like an long-time average.

Example

Let a random variable X to be a number of the rolled die. Then

$$
\begin{aligned}
E(X) &= 1\cdot\frac{1}{6}+2\cdot\frac{1}{6}+3\cdot\frac{1}{6}+4\cdot\frac{1}{6}+5\cdot\frac{1}{6}+6\cdot\frac{1}{6}=3.5 \\
E(X^2) &= 1^2\cdot\frac{1}{6}+2^2\cdot\frac{1}{6}+3^2\cdot\frac{1}{6}+4^2\cdot\frac{1}{6}+5^2\cdot\frac{1}{6}+6^2\cdot\frac{1}{6}=15.16667 \\
E\left(\frac{1}{X}\right) &= 1/1\cdot\frac{1}{6}+1/2\cdot\frac{1}{6}+1/3\cdot\frac{1}{6}+1/4\cdot\frac{1}{6}+1/5\cdot\frac{1}{6}+1/6\cdot\frac{1}{6}=0.40833
\end{aligned}
$$

Note that last two are not equal to $\big(E(X)\big)^2$ and $1/E(x)$.

$$
\begin{aligned}
E(3X+5) &= (3\cdot 1+5)\cdot\frac{1}{6}+(3\cdot 2+5)\cdot\frac{1}{6}+(3\cdot 3+5)\cdot\frac{1}{6} \\
&\quad +(3\cdot 4+5)\cdot\frac{1}{6}+(3\cdot 5+5)\cdot\frac{1}{6}+(3\cdot 6+5)\cdot\frac{1}{6} \\
&= 15.5
\end{aligned}
$$

Note that it is equal to $3E(X)+5$. $\square$

Example: pooled blood testing Each blood sample has .1 change of testing positive. New procedure called pooled testing combines 10 blood sample before testing. If comes back negative, no further test is done. If comes back positive, then 10 more test must be done using individual samples. What is the long-run average of the test number in the new scheme?

Variance

- Variance of a random variable X is defined as

$$V(X) = \sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot p(x_i) = E\Big[\big(X - \mu\big)^2\Big]. = E(X^2) - \Big(E(X)\Big)^2$$

  Note that $\mu = E(X)$. Standard Deviation of X is defined as $\sigma = \sqrt{\sigma^2}$.

- If a and b are constants, then

$$V(aX + b) = a^2 V(X)$$

Example

19 cards in a box.

#	1	2	3	4	5
	3	7	2	6	1
p(x)	3/19	7/19	2/19	6/19	1/19
F(x)	3/19	10/19	12/19	18/19	19/19

$$\begin{aligned} V(X^2) &= (1-2.74)^2 \cdot \frac{3}{19} + (2-2.74)^2 \cdot \frac{7}{19} + (3-2.74)^2 \cdot \frac{2}{19} + \\ &\qquad (4-2.74)^2 \cdot \frac{6}{19} + (5-2.74)^2 \cdot \frac{1}{19} = 1.44 \end{aligned}$$

You could have gotten this by the alternative formula;

$$V(X) = E(X^2) - (E(X))^2 = 8.95 - (2.47)^2 = 1.44.$$

$\square$

Alternative formula for variance can be obtained as

$$\begin{aligned} V(X) &= E\Big[(x-\mu)^2\Big] \\ &= E\Big[x^2 - 2\mu x + \mu^2\Big] \\ &= E[x^2] + E[-2\mu x] + E[\mu^2] \qquad \text{Because of the linearity.} \end{aligned}$$

But $\mu = E(X)$, and they are just a number. Taking E again doesn't do anything. So,

$$\begin{aligned} V(x) &= E[x^2] - 2\mu E[x] + \mu^2 \\ &= E[x^2] - 2\mu^2 + \mu^2 \\ &= E[x^2] - \mu^2 \end{aligned}$$

Similarly to the discrete case, we have

$$V(aX+b) = a^2 V(X)$$

2.2 Popular Discrete RVs

If you have a coin which has probability p of landing head up when tossed, then,

Binomial(n, p)	Number of heads in n tosses
Negative Binomal(r, p)	Number of tails until you get r heads

Other type of discrete random variables includes:

Hypergeometric(n, m, N)	Number of red balls if n balls selected from an urn with N balls which includes m red.
Poisson(λ)	Rare event with rate λ per unit time.

2.2.1 Binomial

Analogy: X is Number of heads in n tosses of a coin. $P(H) = p$ for each toss. [top]

$X \sim Bin(n, p)$

$$\begin{aligned}
\text{pmf}: \quad p(x) &= P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad \text{for } x = 0, 1, 2, \ldots n. \\
\text{CDF}: \quad F(x) &= P(X \leq x) = \sum_{k=0}^{x} p(x) \\
\text{mean}: \quad E(X) &= np \\
\text{var}: \quad V(X) &= np(1-p) \\
\text{MGF}: \quad M(t) &= \left[e^t p + (1-p) \right]^n
\end{aligned}$$

Called Bernoulli Distribution if $r = 1$.

```
dbinom(2, n,  p)       #pmf at x=2
pbinom(2, n,  p)       #CDF at x=2
pbinom(.5, n, p)       #Inv CDF at q=.5
rbinom(1000, n, p)     # random sample of size 1000
```

2.2.2 Negative Binomial

Analogy: Number of **tails** until you get r heads. [top]

$X \sim NegBin(r, p)$

$$
\begin{aligned}
\text{pmf}: \quad p(x) &= \binom{x+r-1}{r-1}(1-p)^x p^r \quad \text{for } x = 0, 1, 2, \ldots \\
\text{CDF}: \quad F(x) &= P(X \le x) = \sum_{k=0}^{x} p(x) \\
\text{mean}: \quad E(X) &= \frac{r(1-p)}{p} \\
\text{var}: \quad V(X) &= \frac{r(1-p)}{p^2} \\
\text{MGF}: \quad M(t) &= [\frac{p}{1-(1-p)e^t}]^r
\end{aligned}
$$

Called Geometric Distribution if $r = 1$.

```
dnbinom(2, r,  p)       #pmf at x=2
pnbinom(2, r,  p)       #CDF at x=2
pnbinom(.5, r,  p)      #Inv CDF at q=.5
rnbinom(1000, r, p)     # random sample of size 1000
```

Negative Binomial (Flips ver.)

Analogy: Number of **flips** until you get r heads. [top]

$X \sim NegBin(r, p)$

$$\begin{aligned}
\text{pmf}: \quad p(x) &= \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad \text{for } x = r, r+1, r+2, \ldots \\
\text{CDF}: \quad F(x) &= P(X \leq x) = \sum_{k=0}^{x} p(x) \\
\text{mean}: \quad E(X) &= \frac{r}{p} \\
\text{var}: \quad V(X) &= \frac{r(1-p)}{p^2} \\
\text{MGF}: \quad M(t) &= [\frac{pe^t}{1-(1-p)e^t}]^r
\end{aligned}$$

Called Geometric Distribution if $r = 1$.

```
dnbinom( 10-r, r,  p)       #pmf at x=10  (In R, X=# of tails)
pnbinom( 10-r, r,  p)       #CDF at x=10
pnbinom(.5, r,  p)      #Inv CDF at q=.5
rnbinom(1000, r, p)     # random sample of size 1000
```

2.2.3 Hypergeometic

Analogy: N balls in an urn, of which m are red. Pick n at once. X= number of red balls. [top]
$X \sim HG(n, m, N)$

$$\begin{aligned}
\text{pmf}: \quad p(x) &= p(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \qquad ** \\
\text{CDF}: \quad F(x) &= P(X \le x) = \sum_{k=0}^{x} p(x) \\
\text{mean}: \quad E(X) &= np \\
\text{var}: \quad V(X) &= \left(\frac{N-n}{N-1}\right) np(1-p) \\
\text{MGF}: \quad M(t) &= DoesNotExist
\end{aligned}$$

where $p = m/N$.

** for $\max(0, n-N+m) \le x \le \min(n, m)$, and 0 otherwise.

```
dhyper(2, m, N-m, n)       #pmf at x=2
phyper(2, m, N-m, n)       #CDF at x=2
phyper(.5, m, N-m, n)      #Inv CDF at q=.5
rhyper(1000, m, N-m, n)     # random sample of size 1000
```

2.2.4 Poisson

Analogy: events with rate λ per unit time. [top]

$X \sim Poi(n, m, N)$

$$\begin{aligned}
\text{pmf}: \quad p(x) &= p(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ for } x = 0, 1, 2, \ldots \\
\text{CDF}: \quad F(x) &= P(X \leq x) = \sum_{k=0}^{x} p(x) \\
\text{mean}: \quad E(X) &= \lambda \\
\text{var}: \quad V(X) &= \lambda \\
\text{MGF}: \quad M(t) &= \exp\{\lambda(e^t - 1)\}
\end{aligned}$$

```
dpois(2, lambda)       #pmf at x=2
ppois(2, lambda)       #CDF at x=2
qpois(.5, lambda)      #Inv CDF at q=.5
rpois(1000, lambda)     # random sample of size 1000
```

Examples

- Multiple Choice question
- Psycic
- Free throw
- Number of tornados

Poisson as a limit If we let $n \to \infty$, $p \to 0$, in such a way that $np \to \lambda$, then the pmf

$$\text{Binomial}(n, p) \to \text{Poisson}(x; \lambda).$$

2.3 Moment Generating Function

- Defined as

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

- Generates moments ($E[X^k]$) as its derivatives at $t = 0$.

$$\begin{aligned} M'_X(0) &= E(X) \\ M''_X(0) &= E(X^2) \\ M'''_X(0) &= E(X^3) \\ &\vdots \end{aligned}$$

Notes

- It may not exist for some r.v.
- mgf only need to exist in an open neighbourhood around 0.
- If they do exist, then mgf specifies cdf.
- Characteristic Function $E(e^{itX})$ always exsits, and specifies cdf.
- Very useful in showing the distribution for sum of independent r.v.
  If $Y = X_1 + X_2$, then

  $$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t(X_1+X_2)}\right] = E\left[e^{t(X_1)}e^{t(X_2)}\right]$$

  If X_1 and X_2 are independent, then

  $$M_Y(t) = E\left[e^{t(X_1)}\right]E\left[e^{t(X_2)}\right] = M_{X_1}(t)M_{X_2}(t)$$

Example:

- Suppose r.v. X has mgf $M_X(t) = 1/(1-t)$. Calculate its mean and variance.
- Obtain pdf of X.

Example:

Suppose Y has MGF $M(t) = e^{3.2(e^t - 1)}$. What is the variance of Y?

Example: K 2-1

Let X_1, X_2, X_3 be a random sample from a discrete distribution with probability function

$$p(x) = \begin{cases} 1/3 & \text{for } x = 0 \\ 2/3 & \text{for } x = 1 \\ & \text{otherwise} \end{cases}$$

Determine the moment generating function of $Y = X_1 + X_2 + X_3$.

2.4 Important Inequalities

[top]

Markov Inequality

- If $E(X)$ exists for **positive** r.v. X, then for every positive constant c,

$$\boxed{P(X \geq c) \leq \frac{E(X)}{c}}$$

Proof: Let $I(\cdot)$ be a indicator function. That is

$$I\{A\} = \begin{cases} 1 & \text{if A is true} \\ 0 & \text{if A is false} \end{cases}$$

Now observe that

$$cI(X \geq c) \leq X.$$

Taking expectation of both side will generate Markov inequality.

$$cP(X \geq c) \leq E(X).$$

$\blacksquare$

Chebyshev Inequality (Tchebysheff)

- Supppose r.v. X has finite variance. Then, for every positive constant k,

$$\boxed{P\Big(|X-\mu| \geq k\sigma\Big) \leq \frac{1}{k^2}}$$

Proof: Use Markov inequality with $X = (X-\mu)^2$, then,

$$P((X-\mu)^2 \geq c) \leq \frac{E\big[(X-\mu)^2\big]}{c}.$$

Now let $c = k^2\sigma^2$. We get

$$\begin{aligned} P((X-\mu)^2 \geq k^2\sigma^2) &\leq \frac{E\big[(X-\mu)^2\big]}{k^2\sigma^2}, \\ P(|X-\mu| \geq k\sigma) &\leq \frac{1}{k^2} \end{aligned}$$

$\blacksquare$

Existance of lower Moments

- If $E[X^m]$ exists, then $E[X^k]$ exists for all positive interger k, m such that $k \leq m$

Example

- Let X be a normal random variable. What is the probability that X is more than 2 σ away from its mean? What does Chebyshev say about this?
- Suppose you have r.v. Y with mean 11 and variance 9. What does Chebyshev say about $P(6 < Y < 16)$?