CH 3 Discrete Random Variables and Their Probability Distributions

mrevanisworking

August 10, 2020

0.1 Basic Definition

0.1.1 Discrete Random Variable (DRV)

a RV Y is said to be discrete if it can assume only a finite or countably infinite num of distinct vals
collection of probabilities is the probability distribution

0.2 Probability Distribution for a DRV

uppercase letter such as Y for RV, and lowercase such as y for particular value of that RV.
$Y = y$ means: the set of all points in S assigned to the value y by the RV Y.

0.2.1 Probability of RV Y

P that Y takes on the value y, $P(Y = y)$ is defined as the sum of the P of all SP in S that are assigned the value y. sometimes noted as $p(y)$
called the probability function for Y

0.2.2 Definition of Probability Distribution

the PD for a discrete var Y can be repr by a formula, table, graph that provides $p(y) = P(Y = y)$ for all y
Any value y not explicitly assigned a probability is understood to be 0

0.2.3 Sum of 1 Theorem

For any discrete PD:
1. $0 \leq p(y) \leq 1$ for all y
2. $\sum_y p(y) = 1$, where the summation is over all vals of y with NZ P

0.3 Expected Value of a RV or a Fxn of RV

0.3.1 Expected Value of DRV

let Y be a DRV w/ P fxn $p(y)$, then the EV of is:

$$E(Y) = \sum_y yp(y)$$

if the sum is convergent.
if $p(y)$ is accurate for the population FD, then $E(Y) = \mu$ is the population mean

0.3.2 HELP Theorem

let Y be DRV w/ P fxn $p(y)$ and $g(Y)$ be a real val Fxn of Y. Then the EV of $g(Y)$ is:

$$E[g(Y)] = \sum_{ally} g(y)p(y)$$

0.3.3 Parameter of $p(y)$ Definition of Variance and SD

if Y is a RV w/ mean $E(Y) = \mu$, the variance of a RV Y is defined to be the EV of $(Y - \mu)^2$:

$$V(Y) = E[(Y - \mu)^2]$$

the SD of Y is the positive square root of $V(Y)$
If $p(y)$ is accurate then $E(Y) = \mu, V(Y) = \sigma^2$ and σ for the population SD

0.3.4 EV constant Theorem

Let Y be DRV w/ PF $p(y)$ then $E(c) = c$

0.3.5 EV Scalar Multiplication Theorem

Let Y be DRV w/ PF $p(y), g(Y)$ be a fxn of Y, c constant, then

$$E[cg(Y)] = cE[g(Y)]$$

0.3.6 EV Addition Theorem

Let Y be DRV w/ PF $p(y), g1tok(Y)$ be k fxns of Y. Then

$$E[g1tok(Y)] = E[g_1(Y)] + \cdots + E[g_k(Y)]$$

0.3.7 EV ffset Theorem

Let Y be DRV w/ PF $p(y)$ and mean $E(Y) = \mu$, then

$$V(Y) = \sigma^2 = E[(Y - \mu)^2] = E(Y^2) - \mu^2$$

0.4 Binomial Probability Distribution

0.4.1 Definition of Binomial Experiment

A binomial experiment:
1. the experiment consists of a fixed num, n, of identical trials
2. Each trial results in one of two outcomes: sucess S or failure F
3. The P of success on a single trial is equal to some value p and remains the same form from trial to trial. The P of a failure is equal to $q = (1 - p)$
4. trials are independent
5. the RV of interest is Y, the num of successes observed during n trials.

(p is probability of success, q is probability of failure)

0.4.2 Binomial Distribution Theorem

An RV Y has a binomial distribution based on n trials w/ success P p IFF

$$p(y) = \binom{n}{y} p^y q^{n-y}$$

where y0ton and $0 \leq p \leq 1$
see 103, FIG3.4
see binomial expansion:

$$\sum_{y=0}^{n} \binom{n}{y} p^y q^{n-y} = (q + p)^n = 1$$

0.4.3 Mean and Varaince of a Binomial Random Variable

Let y be a binomial RV based on n trials and success P p. then

$$\mu = E(Y) = np \text{ and } \sigma = V(Y) = npq$$

Common tricks: $\sum p(y) = 1$
$E(Y^2) = ?E[Y(Y-1)]$
see method of maximum likelihood in CH9

0.5 The Geometric Probability Distribution

Events denoted by on witch trial is the first success
trials are independent

0.5.1 Geometric Probability Distribution

an RV Y is said to have a geometric probability distribution IFF

$$p(y) = q^{y-1}p$$

for y1toinf and $0 \leq p \leq 1$
often used to model distributions of lengths of waiting times

0.5.2 Mean and Variance of Geometric PD

$$\mu = E(Y) = \frac{1}{p} \text{ and } \sigma^2 = V(Y) = \frac{1-p}{p^2}$$

0.6 Negative Binomial Probability Distribution

R2:
knowing the number of the trial on which the rth success occurs

0.6.1 Negative BPD

a RV Y has a NBPD IFF

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r}$$

$y = r, r+1, r+2 ...,\ 0 \leq p \leq 1$

0.6.2 Mean, Variance of NBPD

$$\mu = E(Y) = \frac{r}{p} \text{ and } \sigma^2 = V(Y) = \frac{r(1-p)}{p^2}$$

0.7 The Hypergeometric Probability Distribution

0.7.1 Hypergeometric Probability Distribution

RV Y has a HGPD IFF

$$p(y)\frac{\binom{r}{y}}{\binom{N-r}{n-y}}\binom{N}{n}$$

where y is an int 1ton: $y \leq r, n-y \leq N-r$

0.7.2 Mean and Variance of HGPD

Y is RV w/ HGPD

$$\mu = E(Y) = \frac{nr}{N} \text{ and } \sigma^2 = V(Y) = n(\frac{r}{N})(\frac{N-r}{N})(\frac{N-n}{N-1})$$

see 127 for relation to BPD
Y is approx a BPD when N is large and n is realtively small.
the limit as N goes to infinity is a BPD

0.8 The Poisson Probability Distribution HELP

131 derived the Possion distribution:

$$\lim_{n\to\infty}(1-\frac{\lambda}{n})^n = e^{-\lambda}$$

obtains

$$p(y) = \frac{\lambda^y}{y!}e^{-\lambda}$$

for small p and large n, the binomial counterpart can be used.
good distribution for rare occasions (accidents)

0.8.1 Poisson Probability Distribution

RV Y is PPD IFF

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}$$

y0toinf, $\lambda > 0$

0.8.2 Mean, Varaince for PPD

$$\mu = E(Y) = \lambda \text{ and } \sigma^2 = V(Y) = \lambda$$

0.9 Moments and Moment-Generating Functions

μ, σ do not provide a unique characterization of the distribution of Y
MGF can be used to find moments associated with RV
MGF can be used to establish equivalence of two PD.

0.9.1 Definition of Moment about the origin

the kth moment of a RV Y taken about the origin is defined to be $E(Y^k)$ and is denoted by μ'_k, where $\mu'_2 = E(Y^2)$ is employed in T3.6 for finding σ^2

0.9.2 Definition of Moment about its mean

the kth moment of a RV Y taken about its mean, (kth central moment of Y), is $E[(Y - \mu)^k]$ denoted by μ_k, where $\sigma^2 = \mu_2$

0.9.3 Definition of Moment Generation Function

moment-generating fxn m(t) for a RV Y is defined to be $m(t) = E(e^{tY})$. a MGF for Y exists if there exists a positive constant b such that $m(t)$ is finite for $|t| \leq b$

0.9.4 kth Derivative of MGF Theorem

if m(t) exists, then for any positive int k:

$$\frac{d^k m(t)}{dt^k}\Big]_{t=0} = m^{(k)}(0) = \mu'_k$$

Meaning: kth derivative of $m(t)$ wrt t and then set $t = 0$, result is μ'_k

0.10 Probability-Generating Functions HELP R2

previous PD were all DRV Y that takes positive int vals (binomial, geometric, hypergeometric, Poisson)

0.10.1 Definition of Probability-Generating Funxtion

Y be an int val RV for which $P(Y = i) = p_i$ where i0toinf. The PGF P(t) for Y is:

$$P(t) = E(t^Y)p_0 + p_1 t + p_2 t^2 + \cdots = \sum_{i=0}^{\infty} p_i t^i$$

for all t such that t is finite

0.11 Summary

see 150