March 17, 2025

Last time

Ch. 4 Random Variables:

- X: function from sample space to a set T (T = usually real numbers)

- Probability mass function

$$p(a) = P\{X = a\}$$

- Cumulative distribution function

$$F(x) = P\{X \leq x\}$$

Today: Expected Value (4.3)

EXAMPLE (see also Ch. 2 Prob #35 for similar)

An urn contains 3 red balls
2 blue balls
1 green ball

Three balls are drawn without replacement.

Let X be the number of red balls in the sample.

Note: X is a random variable that can take on values 0, 1, 2, 3.

Zero red

$$p(0) = P\{X=0\} = \frac{\overset{\text{Red}}{\binom{3}{0}}\overset{\text{Blue}}{\binom{2}{2}}\overset{\text{Green}}{\binom{1}{1}}}{\binom{6}{3}} = \frac{1\cdot1\cdot1}{20} = \frac{1}{20}$$

probability mass function

probability that X takes on the value 0

One red

$$p(1) = P\{X=1\} = \frac{\overset{R}{\binom{3}{1}}\overset{B}{\binom{2}{2}}\overset{G}{\binom{1}{0}}}{\binom{6}{3}} + \frac{\overset{R}{\binom{3}{1}}\overset{B}{\binom{2}{1}}\overset{G}{\binom{1}{1}}}{\binom{6}{3}} = \frac{3+3\cdot2}{20} = \frac{9}{20}$$

Two red

$$p(2) = P\{X=2\} = \frac{\overset{R}{\binom{3}{2}}\overset{B}{\binom{2}{1}}\overset{G}{\binom{1}{0}}}{\binom{6}{3}} + \frac{\overset{R}{\binom{3}{2}}\overset{B}{\binom{2}{0}}\overset{G}{\binom{1}{1}}}{\binom{6}{3}} = \frac{3\cdot2+3}{20} = \frac{9}{20}$$

Three red

$$p(3) = P\{X=3\} = \frac{\overset{R}{\binom{3}{3}}\overset{B}{\binom{2}{0}}\overset{G}{\binom{1}{0}}}{\binom{6}{3}} = \frac{1\cdot1\cdot1}{20} = \frac{1}{20}$$

$$F(x) = \begin{cases} 0 & x<0 \\ 1/20 & 0\le x<1 \\ 10/20 & 1\le x<2 \\ 19/20 & 2\le x<3 \\ 1 & 3\le x \end{cases} = P\{X\le x\}$$

← cumulative distribution function

as a check, write out all cases...

(97.1)

$\binom{6}{3} = 20$

		R	R	R	B	B	G
$\frac{1}{20}$	3R	x	x	x			
$\frac{9}{20}$	2R	x	x		x		
		x	x			x	
		x	x				x
		x		x	x		
		x		x		x	
		x		x			x
		x		x			
			x	x	x		
			x	x		x	
			x	x			x
$\frac{9}{20}$	1R	x			x	x	
		x			x		x
		x				x	x
			x		x	x	
			x		x		x
			x			x	x
				x	x	x	
				x	x		x
				x		x	x
$\frac{1}{20}$	0R				x	x	x

20 cases here...

repeat previous example <u>with</u> replacement.

<u>zero red</u>

see Ch. 1 discussion on Multinomial Coefficients (p. 9-10)

$$p(0) = P\{X=0\} = \sum_{i=0}^{3} \left(\frac{3}{6}\right)^0 \left(\frac{2}{6}\right)^i \left(\frac{1}{6}\right)^{3-i} \cdot \underbrace{\frac{3!}{0!\,i!\,(3-i)!}}$$

probability mass function

\# of ways to get 0 red, i blue, $3-i$ green

$$= \sum_{i=0}^{3} \frac{2^i}{6^i \cdot 6^{3-i}} \cdot \frac{3!}{i!\,(3-i)!}$$

$$= \sum_{i=0}^{3} \frac{2^i}{6^3} \cdot \binom{3}{i} = \frac{1}{6^3}\left[1 + 2\cdot 3 + 4\cdot 3 + 8\right] = \frac{27}{6^3}$$

<u>one red</u>

$$p(1) = P\{X=1\} = \sum_{i=0}^{2} \left(\frac{3}{6}\right)\left(\frac{2}{6}\right)^i \left(\frac{1}{6}\right)^{2-i} \frac{3!}{1!\,i!\,(2-i)!}$$

$$= \frac{3}{6^3} \sum_{i=0}^{2} 2^i \frac{3!}{i!\,(2-i)!} = \frac{3}{6^3}\left[1\cdot 3 + 2\cdot 6 + 4\cdot 3\right]$$

$$= \frac{3(27)}{6^3} = \frac{81}{6^3}$$

<u>two red</u>

$$p(2) = P\{X=2\} = \sum_{i=0}^{1} \left(\frac{3}{6}\right)^2 \left(\frac{2}{6}\right)^i \left(\frac{1}{6}\right)^{1-i} \frac{3!}{2!\,i!\,(1-i)!}$$

$$= \frac{9}{6^3} \sum_{i=0}^{1} 2^i \frac{3!}{2!\,i!\,(1-i)!} = \frac{9}{6^3}\left[1\cdot 3 + 2\cdot 3\right] = \frac{81}{6^3}$$

<u>three red</u>

$$p(3) = P\{X=3\} = \left(\frac{3}{6}\right)^3 \left(\frac{2}{6}\right)^0 \left(\frac{1}{6}\right)^0 \frac{3!}{3!\,0!\,0!} = \frac{27}{6^3}$$

Note: $p(0) + p(1) + p(2) + p(3) = 1$

$$F(x) = \frac{1}{216}\begin{cases} 0 & x<0 \\ 27 & 0 \le x < 1 \\ 27+81 & 1 \le x < 2 \\ 27+81+81 & 2 \le x < 3 \\ 216 & 3 \le x \end{cases}$$

Cumulative distribution function.

$$F(x) = P\{X \le x\}$$

comment on General situation

The probability of choosing

R red	balls from	N_R red balls
B blue		N_B blue "
G green		N_G green "

with replacement

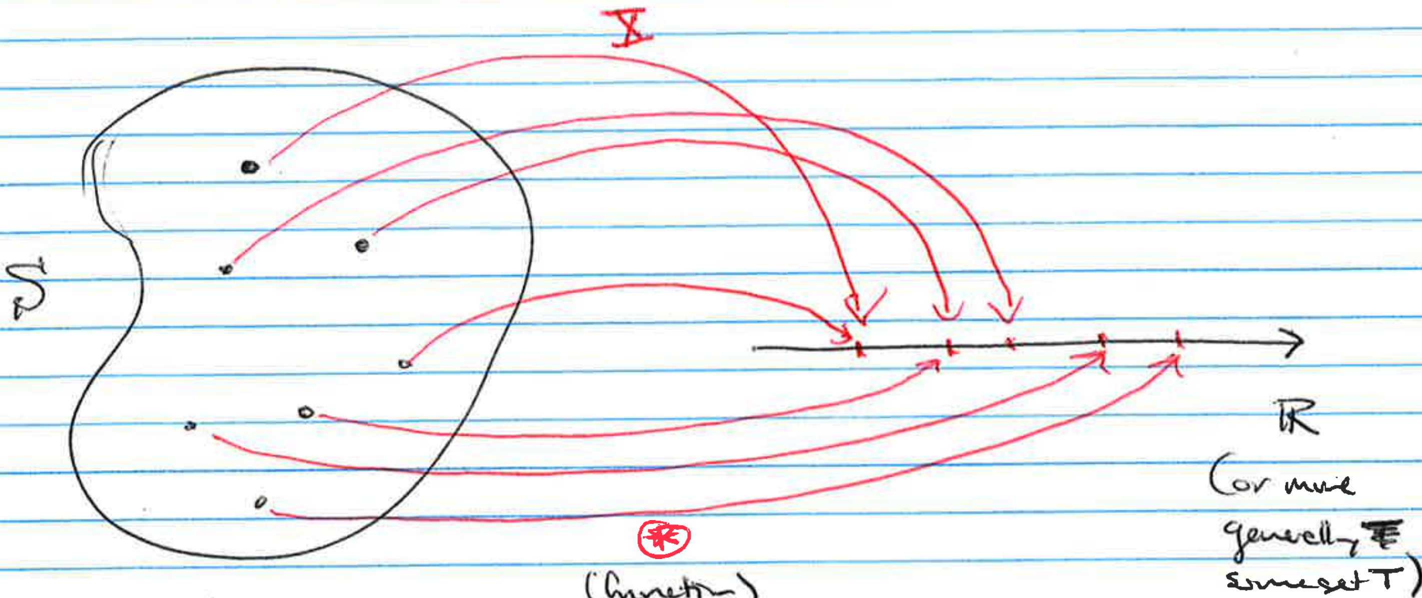
(Note: $R \le N_R$, $B \le N_B$, $G \le N_G$ would be required in the case w/o replacement)

is

$$P = \left(\frac{N_R}{N_R+N_B+N_G}\right)^R \left(\frac{N_B}{N_R+N_B+N_G}\right)^B \left(\frac{N_G}{N_R+N_B+N_G}\right)^G \cdot \underbrace{\frac{(R+B+G)!}{R!\,B!\,G!}}$$

multinomial coefficient representing the number of ways to pick R red, B blue, and G green balls from the $R+B+G$ balls.

Ch.4 continued - Random Variables (Ch. 4.1-4.2 recap)



(function)
random variable

EXAMPLES

outcomes of flipping coin N times	$\xrightarrow{X}$	Number of heads that appear
outcomes of 5 card poker hands	$\xrightarrow{X}$	Some numerical ranking of the hands (i.e. which hand wins)
outcome of a spinner	$\xrightarrow{X}$	- a number between 0 and 2π OR - a discrete set $\{0,1,2,3\}$ ⋮
rolls of ~~rolling~~ two fair dice	$\xrightarrow{X}$	e.g. Sum of the two numbers $\{2,3,\ldots,12\}$ ⋮

A random variable X is a function from S to $\mathbb{R}$ (or more generally T = some set)

- Probability of X taking on a particular value

$$P\{X=i\} = P(\underbrace{\{s \in S : X(s)=i\}}_{\text{could be multiple values of } s \text{ where } X(s)=i.})$$

- The <u>probability mass function</u> p of X is defined by

$$p(i) = P\{X=i\}$$

- The <u>cumulative distribution function</u> F of X is defined by

$$F(x) = P\{X \leq x\} \qquad -\infty < x < \infty$$

<u>4.3 Expected Value</u>

<u>Def:</u> If X is a discrete random variable having probability mass function $p(x)$, then the <u>expected value of X</u> (or expectation of X), denoted by $E[X]$, is defined by

$$E[X] = \sum_{\substack{\text{all } x \\ \text{where} \\ p(x)>0}} x\,p(x)$$

$E[X]$ is also known as "mean of X" or "first moment of X"

EX

~~Toss~~ Flip a fair coin $2N$ times. Let X be a random

variable that denotes the number of heads that occur.

N=1 case (2 flips)

- possible outcomes

assigned values for X

HH	$X(HH) = 2$	$P\{X = 2\} = \frac{1}{4} = p(2)$
HT	$X(HT) = 1$	$P\{X = 1\} = \frac{2}{4} = p(1)$
TH	$X(TH) = 1$	
TT	$X(TT) = 0$	$P\{X = 0\} = 1/4 = p(0)$

- the expected value of X is

$$E[X] = \sum x p(x) = \underset{x=2}{2} \cdot \underset{p(2)}{\left(\frac{1}{4}\right)} + \underset{x=1}{1} \cdot \underset{p(1)}{\left(\frac{2}{4}\right)} + \underset{x=0}{0} \cdot \underset{p(0)}{\left(\frac{1}{4}\right)}$$

$$= \frac{2}{4} + \frac{2}{4} = \frac{4}{4} = \boxed{1}$$

N=2 case (4 flips)

- possible outcomes

HHHH	$X(HHHH)=4$	$P\{X=4\}=\frac{1}{2^4}=p(4)$
HHHT	$X(HHHT)=3$	
HHTH	$\vdots$	$P\{X=3\}=\frac{\binom{4}{3}}{2^4}=\frac{4}{2^4}=p(3)$
HTHH		
THHH		
HHTT	$X(HHTT)=2$	
$\vdots$	$\vdots$	$P\{X=2\}=\frac{\binom{4}{2}}{2^4}=\frac{6}{2^4}=p(2)$
	$X(HTTT)=1$	$P\{X=1\}=\frac{\binom{4}{1}}{2^4}=\frac{4}{2^4}=p(1)$
	$\vdots$	
	$X(TTTT)=0$	$P\{X=0\}=\frac{\binom{4}{0}}{2^4}=\frac{1}{2^4}=p(0)$

$\underbrace{}_{2^4 \text{ possible outcomes}}$

In general, $P\{X=i\} = \underbrace{\left(\tfrac{1}{2}\right)^i}_{i \text{ Heads}} \underbrace{\left(\tfrac{1}{2}\right)^{4-i}}_{4-i \text{ tails}} \cdot \underbrace{\frac{4!}{i!(4-i)!}}_{=\binom{4}{i}}$

$$= \left(\tfrac{1}{2}\right)^4 \binom{4}{i} = \frac{\binom{4}{i}}{2^4}$$

So

$$E[X] = 4\cdot p(4) + 3\,p(3) + 2\cdot p(2) + 1\cdot p(1) + 0\cdot p(0)$$

$$= \frac{1}{2^4}\left[4\cdot 1 + 3\cdot 4 + 2\cdot 6 + 1\cdot 4 + 0\cdot 1\right]$$

$$= \frac{1}{2^4}[32] = \frac{32}{16} = 2 \qquad E[X]=2$$

General case 2N flips

recall: X = # of heads that occur.

$$P\{X=i\} = \frac{\binom{2N}{i}}{2^{2N}}$$

so
$$E[X] = \sum_{i=0}^{2N} i \frac{\binom{2N}{i}}{2^{2N}} = \frac{1}{2^{2N}} \sum_{i=0}^{2N} \binom{2N}{i} i$$

Recall Binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

- If $x=1,\ y=1,\ n=2N,\ k=i$

$$(1+1)^{2N} = \sum_{i=0}^{2N} \binom{2N}{i} 1^i \cdot 1^{2N-i}$$

$$2^{2N} = \sum_{i=0}^{2N} \binom{2N}{i}$$

*

Observe

$$i\binom{n}{i} = \frac{n!}{(n-i)!\,i!}\,i = \frac{n!}{(n-i)!\,(i-1)!} = \frac{n\cdot(n-1)!}{(n-i)!\,(i-1)!}$$

$$= n\binom{n-1}{i-1}$$

So in our problem

$$\sum_{i=0}^{2N}\binom{2N}{i}i = \sum_{i=1}^{2N}\binom{2N}{i}i = \sum_{i=1}^{2N} 2N\binom{2N-1}{i-1} = 2N\sum_{i=1}^{2N}\binom{2N-1}{i-1}$$

Revisit Binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$$

$x=y=1$

$$2^n = \sum_{k=0}^{n}\binom{n}{k} = \sum_{k=1}^{n+1}\binom{n}{k-1}$$

$m \equiv n+1$

$$= \sum_{k=1}^{m}\binom{m-1}{k-1} = \sum_{k=1}^{2N}\binom{2N-1}{k-1}$$

$m=2N$, or $n=2N-1$

So

$$\sum_{i=0}^{2N}\binom{2N}{i}i = 2N\left[\sum_{i=1}^{2N}\binom{2N-1}{i-1}\right] = 2N\cdot 2^{2N-1}$$

Therefore

$$E[X] = \frac{1}{2^{2N}}\left[\sum_{i=0}^{2N} i\binom{2N}{i}\right] = \frac{2N\cdot 2^{2N-1}}{2^{2N}} = \frac{2N}{2} = N$$

$E[X]=N$ as expected

EX Back to urn problem: (see p. 97 in notes)

Urn with 3 red balls
2 blue balls
1 green ball

Draw 3 balls w/o replacement. (second case w/ replacement)

X = number of red balls selected. (random variable)

What is $E[X]$?

Case w/o replacement

$$E[X] = 0 \cdot \frac{1}{20} + 1 \cdot \frac{9}{20} + 2 \cdot \frac{9}{20} + 3 \cdot \frac{1}{20}$$

$$= \frac{9 + 18 + 3}{20} = \frac{30}{20} = 1.5$$

so $\boxed{E[X] = 1.5}$

note: Range of X is $\{0,1,2,3\}$ so $E[X]$ is not necessarily in the range of X

case with replacement see p. 98 notes

$$E[X] = 0 \cdot \frac{27}{6^3} + 1 \cdot \frac{81}{6^3} + 2 \cdot \frac{81}{6^3} + 3 \cdot \frac{27}{6^3}$$

$$= \frac{1}{6^3}\left[81 + 2 \cdot 81 + 3 \cdot 27\right] = \frac{324}{6^3} = 1.5$$

$$\boxed{E[X] = 1.5}$$

EX

A urn contains 4 red balls, and 5 green balls.

Three balls are chosen w/o replacement.

Let X = # of red balls chosen. Compute $P\{X=i\}$ $i=0,1,2,3$

and $E[X]$.

i	0	1	2	3
$P\{X=i\}$:	$\frac{\binom{4}{0}\binom{5}{3}}{\binom{9}{3}}$	$\frac{\binom{4}{1}\binom{5}{2}}{\binom{9}{3}}$	$\frac{\binom{4}{2}\binom{5}{1}}{\binom{9}{3}}$	$\frac{\binom{4}{3}\binom{5}{0}}{\binom{9}{3}}$
	$=\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$

so

$$E[X] = 0\cdot\frac{10}{84} + 1\cdot\frac{40}{84} + 2\cdot\frac{30}{84} + 3\cdot\frac{4}{84}$$

$$= \frac{40+60+12}{84} = \frac{112}{84} = \boxed{\frac{4}{3}} \qquad E[X]=\frac{4}{3}$$

with replacement:

i	0	1	2	3
$P\{X=i\}$	$\left(\frac{4}{9}\right)^0\left(\frac{5}{9}\right)^3\cdot\underbrace{\frac{3!}{0!3!}}_{\binom{3}{0}}$	$\left(\frac{4}{9}\right)^1\left(\frac{5}{9}\right)^2\underbrace{\frac{3!}{1!2!}}_{\binom{3}{1}}$	$\left(\frac{4}{9}\right)^2\left(\frac{5}{9}\right)^1\underbrace{\frac{3!}{2!1!}}_{\binom{3}{2}}$	$\left(\frac{4}{9}\right)^3\left(\frac{5}{9}\right)^0\cdot\underbrace{\frac{3!}{0!3!}}_{\binom{3}{3}}$
	$=\frac{125}{729}$	$\frac{300}{729}$	$\frac{240}{729}$	$\frac{64}{729}$

$$E[X] = 0\cdot\frac{125}{729} + 1\cdot\frac{300}{729} + 2\cdot\frac{240}{729} + \frac{3\cdot 64}{729} = \frac{972}{729} = \frac{4}{3} \qquad E[X]=\frac{4}{3}$$

EX

An urn contains N_R red balls and N_G green balls

N balls are chosen (a) w/o replacement, (b) with replacement.

Assume $N_R \geq N$ $N_G \geq N$

Let X = random variable = # of red balls chosen.

Compute $P\{X=i\}$ $i=0,1,2,\ldots,N$

and $E[X]$

sample size n, urn w/ N balls, m red $N-m$ green

This is a <u>hypergeometric</u> random variable, which has the form $P\{X=i\} = \dfrac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$

see section 4.8.3

a) w/o replacement

$$\boxed{P\{X=i\} = p(i) = \frac{\binom{N_R}{i}\binom{N_G}{N-i}}{\binom{N_R+N_G}{N}} \quad \text{for } i=0,1,2,\ldots,N}$$

$$E[X] = \sum_{i=0}^{N} i\,p(i)$$

$$= \sum_{i=0}^{N} i \frac{\dfrac{N_R!}{i!(N_R-i)!} \cdot \dfrac{N_G!}{(N-i)!(N_G-N+i)!}}{\dfrac{(N_R+N_G)!}{N!(N_R+N_G-N)!}}$$

$$E[X] = \frac{N!(N_R+N_G-N)!}{(N_R+N_G)!} N_R! N_G! \sum_{i=0}^{N} \frac{i}{i!(N_R-i)!(N-i)!(N_G-N+i)!}$$

107.1.1

$$E[X] = \sum_{i=0}^{N} i \frac{\binom{N_R}{i}\binom{N_G}{N-i}}{\binom{N_R+N_G}{N}}$$

$$= \sum_{i=1}^{N} \frac{i\binom{N_R}{i}\binom{N_G}{N-i}}{\binom{N_R+N_G}{N}}$$

note discussion on hypergeometric random variable - example 8j (p.153).

use identities $\quad i\binom{N_R}{i} = N_R\binom{N_R-1}{i-1}$

and $\quad N\binom{N_R+N_G}{N} = (N_R+N_G)\binom{N_R+N_G-1}{N-1}$

$$E[X] = \sum_{i=1}^{N} \frac{N_R\binom{N_R-1}{i-1}\binom{N_G}{N-i}}{\frac{N_R+N_G}{N}\binom{N_R+N_G-1}{N-1}}$$

$$= N \cdot \frac{N_R}{N_R+N_G} \underbrace{\sum_{i=1}^{N} \frac{\binom{N_R-1}{i-1}\binom{N_G}{N-i}}{\binom{N_R+N_G-1}{N-1}}}_{\substack{\text{let } i=j+1 \\ N=M+1}}$$

$$= N \cdot \frac{N_R}{N_R+N_G} \cdot \underbrace{\boxed{\sum_{j=0}^{M} \frac{\binom{N_R-1}{j}\binom{N_G}{M-j}}{\binom{N_R+N_G-1}{M}}}}_{\substack{\| \\ 1}}$$

$= p(j)$ for probability of selecting j red balls from an urn with N_R-1 red and N_G green balls.

since $\sum p(j)=1$ (prob. density function)

$$\boxed{E[X] = N \cdot \frac{N_R}{N_R+N_G}}$$

Note: $M = N-1 \le N_G - 1$
$M = N-1 \le N_R - 1$

b) with replacement

$$\boxed{P\{X=i\} = p(i) = \left(\frac{N_R}{N_R+N_G}\right)^i \left(\frac{N_G}{N_R+N_G}\right)^{N-i} \cdot \frac{N!}{i!(N-i)!} \quad \text{for } i = 0,1,2,\ldots,N}$$

Note: this has the form $p(i) = \binom{n}{i} p^i (1-p)^{n-i}$ which is a binomial random variable

$\boxed{\text{Section } 4.6}$

$$E[X] = \sum_{i=0}^{N} i\,p(i)$$

$$= \sum_{i=0}^{N} i \left(\frac{N_R}{N_R+N_G}\right)^i \left(\frac{N_G}{N_R+N_G}\right)^{N-i} \frac{N!}{i!\,(N-i)!}$$

$$\cancel{= \frac{N!}{(N_R+N_G)^N} \sum_{i=0}^{N} N_R^{\,i}\, N_G^{\,(N-i)} \frac{i}{i!\,(N-i)!}}$$

$$\cancel{= \frac{N!}{(N_R+N_G)^N} \sum_{i=0}^{N} N_R^{\,i}\, N_G^{\,(N-i)}}$$

use $i\binom{N}{i} = N\binom{N-1}{i-1}$

$$= \sum_{i=0}^{N} \left(\frac{N_R}{N_R+N_G}\right)^i \left(\frac{N_G}{N_R+N_G}\right)^{N-i} \underbrace{i\binom{N}{i}}_{= N\binom{N-1}{i-1}}$$

replace with $i=1$ since $i=0$ case evaluates to zero

$$= N \sum_{i=1}^{N} \left(\frac{N_R}{N_R+N_G}\right)^i \left(\frac{N_G}{N_R+N_G}\right)^{N-i} \binom{N-1}{i-1}$$

let $a = \frac{N_R}{N_R + N_G}$, $b = \frac{N_G}{N_R + N_G}$, note: $a + b = 1$

(107.3)

$$E[X] = N \sum_{i=1}^{N} a^i b^{N-i} \binom{N-1}{i-1}$$

Recall the binomial theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

let $i = j+1$

$$E[X] = N \cdot \sum_{j=0}^{N-1} a^{j+1} b^{N-(j+1)} \binom{N-1}{j}$$

let $M = N-1$

$$E[X] = N \cdot \sum_{j=0}^{M} a^{j+1} b^{M+1-(j+1)} \binom{M}{j}$$

$$= N \sum_{j=0}^{M} a\, a^j b^{M-j} \binom{M}{j}$$

$$= aN \boxed{\sum_{j=0}^{M} a^j b^{M-j} \binom{M}{j}}$$

from binomial thm. this is $(a+b)^M$

$$= aN (a+b)^M$$ but $a+b=1$

so $E[X] = a \cdot N = \boxed{N \cdot \left(\frac{N_R}{N_R + N_G} \right) = E[X]}$ with replacement

For the case

$$N_R \geq N$$
$$N_G \geq N$$

we see that $E[X]$ is the same for both cases — without replacement AND with replacement — both are

$$E[X] = \frac{N_R}{N_R + N_G} \cdot N$$

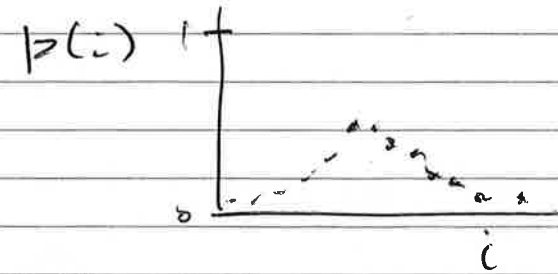
see p. 107.1.1 – without replacement
p. 107.3 – with replacement

Note, however that the probability distribution functions $p(i)$ for the two cases are not the same.

See also Matlab code

two_color_urn_without_replacement.m

that plots



and computes $E[X]$.

The proof of the "without replacement" case relates to hypergeometric random variables (which X is in that case). (see Section 4.8.3 in Ross, Edition #9)

EXAMPLE

Define I as the indicator variable for the event A as

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

(e.g. Flip of coin
Heads/Tails   Heads = 1   A = heads occurs
Tails = 0 A^c = tails occurs)

Note that the expected value of I is

$$E[I] = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A)$$