

Recall

Sampling with (out) replacement

Suppose a village has 5000 people

- 1000 people are below age of 18.

[1] Choose 4 people at "random" from the population.

SRSWOR

- 5000 C₄ = number of 4 people & each group is chosen in equally likely manner.

IP(the age

+

[2]

ch

Problem

Difference from (1)

SRSWR

Recall

Sampling with (out) replacement

Suppose a village has 5000 people

- 1000 people are below age of 18.

[1] Choose 4 people at "random" from the population.

SRSWOR

- 5000 C_4 groups of 4 people & each group is chosen in equally likely manner.

Pr the

4

[2]

Problem
Difference
from (1)

SRSWR

$$P(\text{there are exactly two people below the age of 18 in group of 4 chosen}) = \frac{{}^{1000}C_2 {}^{400}C_2}{{}^{5000}C_4} = 0.153592$$

2 choose four people, one at a time, with replacement (randomly)

Problem
Difference
from ①

SRS WR

- the same person could be chosen twice.

- each choice is a Bernoulli ($\frac{1}{5}$) experiment
(if we are looking for a person below the age of 18)

$P(\text{there are exactly 2 people below the age of 18})$

$= P(2 \text{ successes in Binomial}(4, 1/5) \text{ experiment})$

$$= {}^4C_2 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^2 = 0.1536$$

Note:- Two methods yield very close answers, but are not equal.

- '4' - referred to as sample size
- "500" - " " " population size } two methods yield "similar" answers

When sample size is relatively small compared to population size.

Hypergeometric Distribution (N, r, m)

Consider a population of N people ^{or} [things, ...]

r - out of N have a certain ^{common} characteristic
($N-r$ do not have this characteristic)

m - size of ^{random} sample from N without replacement

$[K]$ - Count the number of elements in the sample that have the characteristic

Count

$$\leq r$$

$$\geq 0$$

$$\geq m - (N - r)$$

$$\leq \min\{m, r\}$$

$$S = \{ \max\{0, m - (N - r)\}, \dots, \min\{m, r\} \}$$

$$P(\{k\}) = \frac{r C_k \cdot N - r C_{m-k}}{N C_m} \quad \text{--- (*)}$$

$$k \in S$$

Such a distribution on S is called the hypergeometric distribution with parameters (N, r, m)

Ex: let $P(\{k\})$ be as in (*)
Show $\sum_{k \in S} P(\{k\}) = 1$

2.3.2 Dependent

- it is possible to have a

Remark Example 2.3.1

$j = 1, 2, 3, 4$ - choose

P2.3.2 Dependent trials

- it is possible to consider hypergeometric Geometric distribution as a set of dependent Bernoulli trials.

* - each trial is like a Bernoulli trial

- probability of success in each trial depends on the previous trials

Remark Example 2.31 - [1000 below age of 18 in a population of 5000]

- choose 4 people one at a time, without replacement
(equally likely names from the population at that trial)

$j = 1, 2, 3, 4$

distribution

$A_j = \{ \text{at the } j\text{th trial, a person below the age of 18 is chosen} \}$

$$B_j = A_j^c$$

$$P(A_1) = \frac{1000}{5000} = \frac{1}{5}, \quad P(B_3 | A_1 \cap A_2) = \frac{4000}{4998}$$

$$P(A_2 | A_1) = \frac{999}{4999}$$

$$P(B_4 | A_1 \cap A_2 \cap B_3) = \frac{3999}{4997}$$

Similar

ren}

let us denote - (s) success = if a person below age of 18 is chosen
 - (f) or failure = otherwise

$$P(\text{success, success, failure, failure}) = P(A_1 \cap A_2 \cap B_3 \cap B_4)$$

①

$$= P(B_4 | A_1 \cap A_2 \cap B_3) \cdot P(B_3 | A_1 \cap A_2) \cdot P(A_2 | A_1) \cdot P(A_1)$$

(Theorem 1.3.8)

-000

+998

$$= \frac{3999}{4997} \cdot \frac{4000}{4998} \cdot \frac{999}{4999} \cdot \frac{1000}{5000}$$

=3999

4997

Similarly one can calculate

Ex 1 - $P(S, f, S, f)$, $P(S, f, f, S)$, $P(f, f, S, S)$, $P(f, S, f, S)$, $P(f, S, S, f)$

② ③ ④ ⑤ ⑥

- we will n
 but

[obs]

P(exactly 2
 from the
 below

- We will notice - that each of these are the same 'product'
but of different fractions

[observe - denominator goes from 5000, 4999, 4998, 4997
numerator will be from 3999, 4000, 999, 1000
but in different orders]

$$P(\text{exactly two are chosen from the population of 5000 below age of 18}) = 6 \cdot \frac{3999 \cdot 4000 \cdot 999 \cdot 1000}{4997 \cdot 4998 \cdot 4999 \cdot 5000} = \frac{1000 \cdot 4000}{5000 \cdot 4}$$

$$S), P(A, S, S, F)$$

Theorem 2.3.2 :- Consider the sample space S
 A hypergeometric (N, r, m) distribution.

$$P(\{k\}) = \frac{{}^r C_k {}^{N-r} C_{m-k}}{{}^N C_m} \quad k \in S$$

Proof
Ex

$$= {}^m C_k \left[\frac{r}{N} \cdot \frac{r-1}{N-1} \cdots \frac{r-(k-1)}{N-(k-1)} \right] \left[\frac{N-r}{N-k} \cdot \frac{N-r-1}{N-k-1} \cdots \frac{N-r-(m-1-k)}{N-(m-1)} \right]$$

$$= {}^m C_k \frac{\prod_{j=0}^{k-1} (r-j)}{\prod_{i=k}^{m-1} (N-i)}$$

Where

Theorem 2.3.3

N, m, r be positive integers

Assume $m < r < N$ k be a positive integer $\in S$ from hypergeometric

$P(k)$ - be the hypergeometric probability

$$\binom{m}{k} p^k (1-p)^{m-k} < P(k) \leq \binom{m}{k} q^k (1-q)^{m-k}$$

where $p = \frac{r}{N}$, $t = \frac{r-k}{N-m}$, $q = \frac{r-k}{N-k}$

Recall

Theorem 2.3.3

$$p = \frac{r}{N}$$

$$t = \frac{r-k}{N-m}$$

$$q = \frac{r-k}{N-k}$$

$$P(\{k\}) = \frac{\binom{r}{k} \binom{N-r}{m-k}}{\binom{N}{m}}$$

$$m \binom{r}{k} q^k (1-t)^{m-k}$$

$$P(\{k\}) \leq$$

$$m \binom{r}{k} p^k (1-t)^{m-k}$$

Sample space
of hypergeometric
(N, r, m)

$$k \in S$$

3

k space
hypergeometric
(N, r, m)

- Gap between Sampling with and without replacement
- N - large & m - small then $p \approx t \approx n$

3 Discrete Random Variables

We have developed many Probability distributions
& in each case we proceed by constructing a sample
space for the same!

Binomial (n, p)
Bernoulli (p)
Geometric (p)
Hypergeometric (N, r, m)
Poisson (λ)
Uniform $\{1, \dots, n\}$

$m-k$
-t)

Inter
func

Suc

Exam

(a) H

(b)

Instead, now we will have one sample space & create several functions whose output relate to the questions in consideration.

Such functions are known as random variables.

Example 3.1.1.

Suppose we flip a coin 3-times.

(a) How many tosses will end up in head?

(b) Which will be the first toss that will be a head?

answer is easy.

but we will use the example to illustrate
how to use functions effectively.

$$S = \{ hhh, hht, hth, thh, tth, tht, htt, ttt \}$$

$$X: S \rightarrow \mathbb{R}$$

$$w \rightarrow X(w) \equiv \# \text{ of heads in } w$$

$$Y: S \rightarrow \mathbb{R} \cup \{\text{none}\}$$

$$w \rightarrow Y(w) \equiv \text{the toss at which first head appears}$$

hhh

hht

hth

thh

tth

tht

htt

ttt

w	$X(w)$	$Y(w)$
hhh	3	1
hht	2	1
hth	2	1
thh	2	2
tth	1	3
tht	1	2
htt	1	1
ttt	0	none

Suppose we want to know

$P(\text{there are exactly 2 heads in 3 tosses})$?

$E = \{\text{there are exactly 2 heads in 3 tosses}\}$

$= \{hht, hth, thh\}$

$= \{w \in S \mid X(w) = 2\}$

$= \bar{x}^{-1}(\{2\})$

$P(hht, hth, thh)$

$\frac{3}{8}$

$P(\text{first head})$

In the above

Using

$$\Rightarrow P(E) = P(X^{-1}(\{2\})) := P(X=2) \quad \text{\textit{create several}} \\ \text{\textit{notation \& well used}} \\ \text{\textit{short form}}$$

$$P(\{hht, hth, thh\})$$

$$\parallel \\ 3/8$$

$$\bullet P(\text{first head occur in third toss}) = P(Y=3) = P(\{tth\}) = 1/8$$

In the above, a single sample space was used to compute different probabilities using functions.

Theorem 3.1.2 Let S be a sample space, with
a probability $\mathbb{P}$ on it. Let $X: S \rightarrow T$
be a function. Then " X generates a probability
 $\mathbb{Q}$ on T ", given by

$$\text{BST, } \mathbb{Q}(B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B)$$

The probability $\mathbb{Q}$ on T is known as the "distribution
of X ", as it describes how X distributes

the $\mathbb{P}$

Proof

(clear

(i)

(ii)

the probability $\mathbb{P}$ from S on T .

Proof: - let $\mathcal{G} \equiv$ ^{collection of all} _{subsets of} T

clearly, $\mathcal{Q}: \mathcal{G} \rightarrow [0,1]$ by definition Axiom 1 applied to $\mathbb{P}$

$$(i) \quad \mathcal{Q}(T) = \mathbb{P}(\bar{X}'(T)) = \mathbb{P}(S) \stackrel{\downarrow}{=} 1$$

(ii) let $\{B_i\}_{i \geq 1}$ be a countable collection of disjoint events in T .

the probability $\mathbb{P}$ from S on T .

Proof: - let $\mathcal{G} \equiv$ ^{collection of all} subsets of T

clearly, $\mathbb{Q}: \mathcal{G} \rightarrow [0,1]$ by definition

Axiom 1 applied to $\mathbb{P}$

$$(i) \quad \mathbb{Q}(T) = \mathbb{P}(\bar{X}'(T)) = \mathbb{P}(S) \stackrel{\downarrow}{=} 1$$

(ii) let $\{B_i\}_{i \geq 1}$ be a countable collection of disjoint events in T .

$$Q\left(\bigcup_{i=1}^{\infty} B_i\right) = P\left(X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right)\right) \quad \text{create several}$$

$$s \in X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right) \Leftrightarrow X(s) \in \bigcup_{i=1}^{\infty} B_i \Leftrightarrow X(s) \in B_i \text{ for some } i \Leftrightarrow s \in \bigcup_{i=1}^{\infty} X^{-1}(B_i)$$

$$X^{-1}\left(\bigcup_{i=1}^{\infty} B_i\right) = \bigcup_{i=1}^{\infty} X^{-1}(B_i) \quad (*)$$

$$E_X: \quad X^{-1}(B_i) \cap X^{-1}(B_j) = \emptyset \quad \forall i \neq j \quad (**)$$

Form (1)

$$\phi(\bigcup_{i=1}^n B_i) = P(\bar{X}^1(\bigcup_{i=1}^n B_i))$$

$$\stackrel{(1)}{=} P(\bigcup_{i=1}^n \bar{X}^1(B_i))$$

Answer (2)

by the P

(2)

$$= \sum_{i=1}^n P(\bar{X}^1(B_i))$$

$$= \sum_{i=1}^n \phi(B_i)$$

(11)

$(I) \subseteq (II) \Rightarrow \mathcal{P}$ satisfies Axiom 1 & 2 in Definition 11.3. $\square$

Proof: - let $\mathcal{G} \equiv$ ^{collection of all} subsets of T

clearly, $\mathcal{P} : \mathcal{G} \rightarrow [0,1]$ by definition Axiom 1 applied to $\mathcal{P}$

$$(i) \quad \mathcal{P}(T) = \mathbb{P}(\bar{X}(T)) = \mathbb{P}(S) \stackrel{\downarrow}{=} 1 - (I)$$

(ii) let $\{B_i\}_{i \geq 1}$ be a countable collection of disjoint events in T .