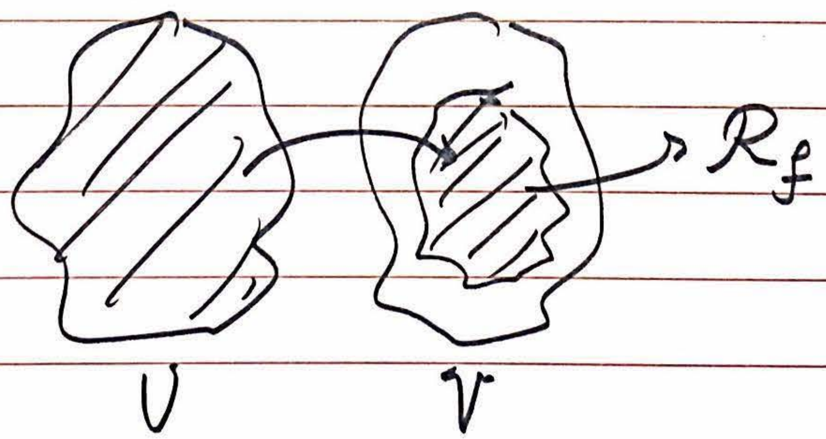
Lesson 7

Discrete Random Variables

Definition (Range):

Assume that $f: U \to V$ is a function. The range of f, R_f, is defined as $\vec{f}(U)$



Def1 (Discrete Random Variable):

A random variable $X: \Omega \to \mathbb{R}$ is called discrete if $\overrightarrow{X}(\Omega) = R_X$ (i.e. the range of X) is <u>countable</u>.

(In other words, X can take discrete values $x_1, x_2, \ldots$)

Def2 (More Precise):

A random variable $X: \Omega \to \mathbb{R}$ is said to be discrete if it takes values in a countable subset of $\mathbb{R}$ with probability 1.

Therefore, there is a countable set $A = \{x_1, x_2, \dots\}$ such that

$$\mathbb{P}(\underbrace{X \in A}_{\text{event}}) = \mathbb{P}_X(A) = 1.$$

This definition does not demand the range of the random variable be countable. There might be a zero probability ~~set~~ event that is mapped to an uncountable set of $\mathbb{R}$.

Remark: We can work with both definitions, but Def 1 is slightly easier, because it focuses on the range of the random variable, i.e. the values that X can take, so we work with the second definition.

Def (Probability Mass Function PMF): If X is a discrete r.v., the function $p_X : \mathbb{R} \to [0,1]$, defined as $p_X(x) = \mathbb{P}(X = x)$ is called the pmf of X.

Interestingly, ~~the~~ a discrete random variable can be completely characterized by its pmf:

For any Borel Set B,

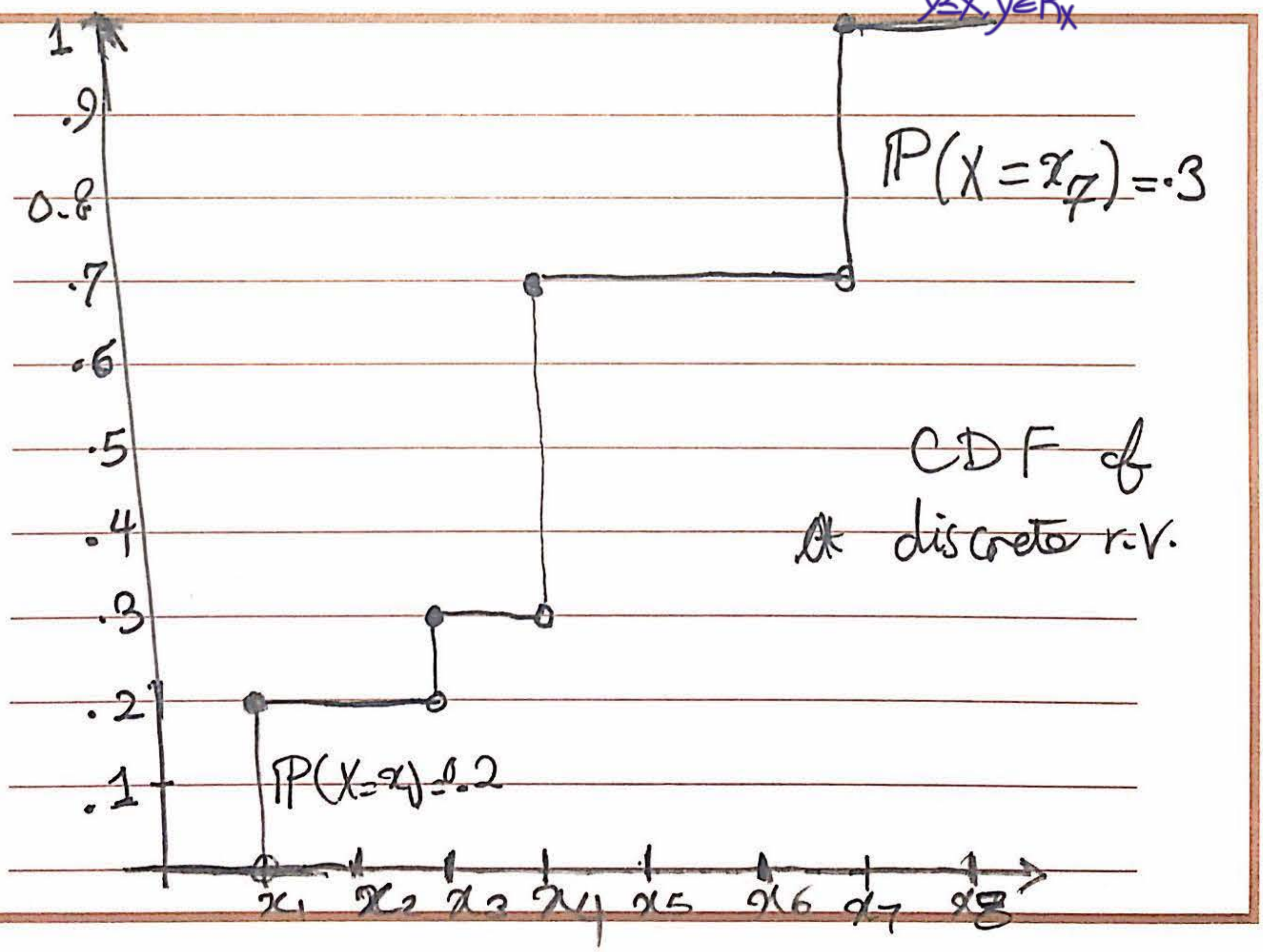
$$\mathbb{P}(X\in B) = \mathbb{P}_X(B) = P(X \in B\cap R_X)$$

$$= \sum_{x\in B\cap R_X} P(X=x) = \sum P_X(x)$$

The cdf of a discrete r.v. is given by:

$$F_X(x) = \mathbb{P}(\underbrace{X \leq x}_{X\in(-\infty,x]}) = \sum_{y\leq x} P_X(y)$$

$$= \sum_{y\leq x,\, y\in R_X} P_X(y)$$



$$P(X=x) = F_X(x) - \lim_{y\to x^-} F_X(y)$$

Obviously $\sum_{x \in \mathbb{R}} P_X(x) = 1$

$$\sum_{x \in \mathbb{R}} P_X(x) = P(X \in \mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$$

or

$$= \sum_{x \in \mathbb{R} \cap \mathbb{R}_X} P(X=x) = \sum_{x \in \mathbb{R}_X} P(X=x) = P(X \in R_X)$$

$$= P(X^{-1}(R_X)) = P(\Omega) = 1$$

Important Discrete Random Variables and Their pmfs

(a) Bernoulli with parameter p where $0 \le p \le 1$

$X \sim Ber(p)$

or $X \stackrel{d}{=} Ber(p)$

$$\begin{cases} P_X(1) = p \\ P_X(0) = 1-p \end{cases}$$

$P(X \in A) = p \quad P(X \in A^c) = 1-p$

A Bernoulli random variable models binary events that have ~~two~~ two possible outcomes: success & failure.

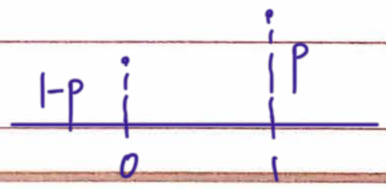
Coin Toss : Success: $\boxed{H}$
Failure: $\boxed{T}$

$\Omega = [0,1]$ success Failure
$x \in A = [0, 1/2]$ $x \in (1/2, 1]$

(I_A is a Bernoulli Random Variable)

$$I_{[0,1/2]} = \begin{cases} 1 & \omega \in [0, 1/2] \\ 0 & \omega \in (1/2, 1] \end{cases}$$

pmf



(b) Discrete Uniform

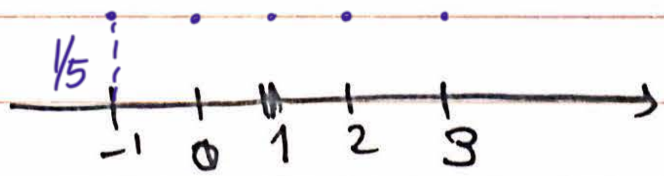
$a, b \in \mathbb{Z} \qquad a < b$

$$P_X(x) = \begin{cases} \frac{1}{b-a+1} & x = a, a+1, \dots, b \\ 0 & \text{otherwise, } x \in \mathbb{Z} \end{cases}$$

$X \sim dU(a,b)$

Example: $a = -1$, $b = 3$

$$P_X(x) = \begin{cases} \frac{1}{3-(-1)+1} = 1/5 & x = -1, \dots, 3 \\ & \text{otherwise} \end{cases}$$



$$\sum_{x=-1}^{3} P_X(x) = 5 \times 1/5 = 1$$

$$0 \le P_X(x) \le 1$$

(c) Binomial with parameters n, p

$n \in \mathbb{N}$, $p \in [0, 1]$

A Binomial r.v. can be ~~seen~~ model as a sequence of n independent Bernoulli trials with success probability p. The Binomial r.v. shows the number of successes.

The pmf of Binomial r.v.'s is:

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$0 \leq X \leq n$ $X \sim \text{Bin}(n, p)$

Justification.

$$p_X(x) = \mathbb{P}(X = x) = \mathbb{P}(E_x)$$

number of success

where

$$E_x = \{\omega \in \Omega \mid \omega \text{ has } x\text{-successes and } n-x\text{-failures}\}$$

From combinatorics, we know

$$|E_x| = \binom{n}{x}$$

Thus $\mathbb{P}(E_x) = \binom{n}{x} p^x (1-p)^{n-x}$

Example: Sara hits 35% of her free throws in basketball games. Yesterday she had 7 free throws.

a) What is the probability that she had between 4 and 7 hits?

$p = 0.35$

$$\sum_{4 \le x \le 7} P_X(x) = \sum_{x=4}^{7} \binom{7}{x} p^x (1-p)^{7-x}$$

(b) What is the probability that she had at most 5 hits?

$$P(X \leq 5) = \sum_{x=0}^{5} \binom{7}{x} p^x (1-p)^{7-x}$$

$$= 1 - P(X > 5)$$

Exercise: The simplest error detection scheme used in data communication is parity-checking. Messages sent consist of characters, each character

consisting of a numbers of 0 or 1 bits. In parity checking a 0 or 1 is appended to the end of the message to make the number of 1's even.

$\underbrace{1110100}_{7\text{ bits}}\ \underbrace{0}_{\text{parity}}$ keeping # of 1's even

$\underbrace{1100100}_{7\text{ bits}}\ \underbrace{1}_{\text{parity}}$ adding 1 to make the # of 1's even

The receiver checks the number

of 1's in every character received and if the result is odd, it signals an error. Suppose that each bit is received correctly with probability $1-\varepsilon_1$ independently of other bits. What is the probability that Assuming that the message is n bits

a) An erratic message is observed in the receiver side?

$\mathbb{P}$(erratic message)

$= \mathbb{P}$(odd # of ones)

$= \mathbb{P}$(odd # of bit-flips)

(b) What is the probability that an erratic message is received and not detected by the parity check?

(d) Geometric with parameter p

We define two kinds of geometric random variables:

i) $X \sim Geo_1(p)$, $p \in (0,1]$

$$p_X(x) = \qquad\qquad x \in \mathbb{N}$$

The Geo_1 random variable represents the number of independent Bernoulli trials needed until the first success is observed

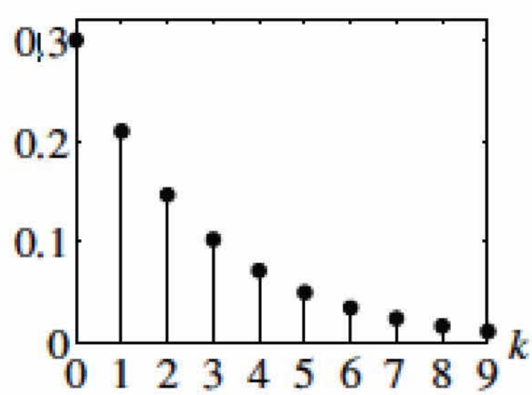
ii) The second type of geometric random variable has the following pmf

$X \sim Geo_0(p)$, $p \in (0,1]$

$p_X(x) =$ $\qquad x \in \{0,1,2,\dots\}$

The pmf of $Geo_0(0.3)$

shown below:



Exercise: Show that when

$X \sim Geo_0(p)$ or $Y \sim Geo_1(p)$,

$$\sum_{x=0}^{\infty} p_X(x) = 1 \text{ and } \sum_{y=1}^{\infty} p_Y(y) = 1,$$

indeed.

Example: A hacker wants to break into a computer file, which is password protected. Assume that the passwords can only be alpha-numeric, i.e. they can only contain a-z, A-Z and 0-9. Assume that the hacker tries possible passwords in a random order. Let X be the number of trials required to break into the file. Determine the pmf

of X: if

a) if unsuccessful passwords

are not eliminated from further

selections

(b) if unsuccessful passwords are

eliminated.

6. a) X has geometric distribution

with $p = \frac{1}{n}$, where $n =$

because it is the number

of independent Bernoulli trials

until the first success

(correct password)

$\mathbb{P}(X = x) =$

(b) Here, the probability of success increases after each failure, so the . . distribution of X is not Geometric. Since the order is random, the correct

password has the same chance to appear in each sequence. Therefore, X has a $\boxed{}$ distribution

$\mathbb{P}(X=x) =$

(e) Poisson with parameter $\lambda > 0$

The Poisson random variable is used to model many different phonomena such as the photoelectric effect, radioactive decay, and computer traffic arriving at a queue for transmission.

$X \sim \text{Pois}(\lambda)$

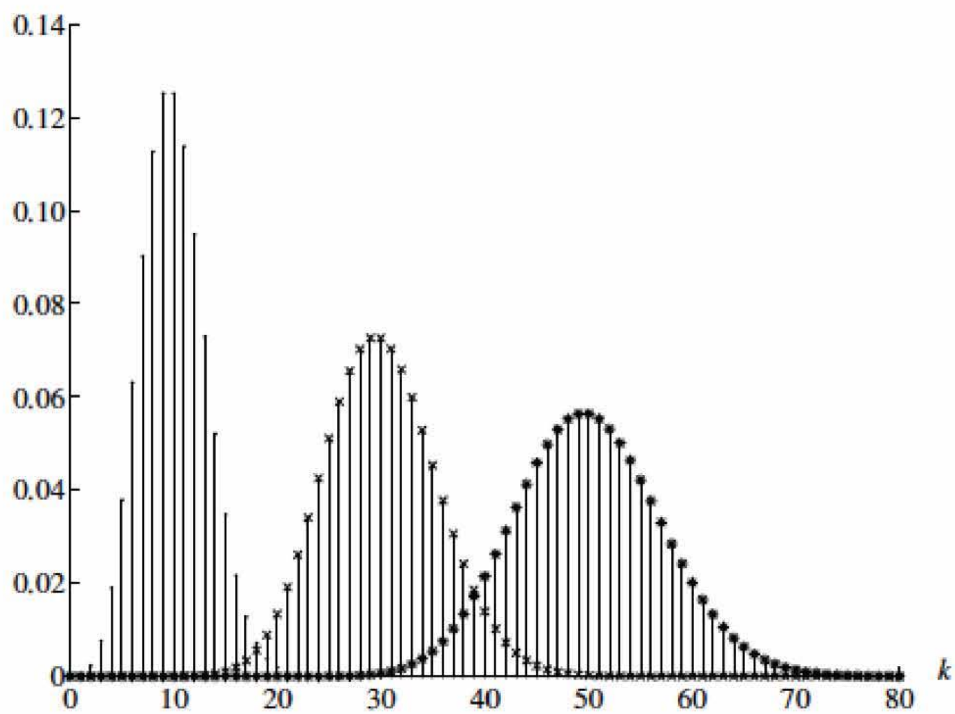
$$P_X(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad x = 0, 1, 2, \dots$$

Recall that

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

Show that the pmf of a Poisson random variable sums to one for any $\lambda > 0$

The pmf of a $Pois(\lambda)$ is shown in the following figure for $\lambda = 10, 30, 50$, from left to right respectively:



The Poisson Distribution was introduced by Abraham de Moivre in 1711 and was further developed by Siméon Denis Poisson in 1837.

The Poisson ______ Random Variable represents the number of events that occur in a fixed time interval (or volume, length etc) assuming that they occur independently of the time since last event arrived, and their average rate in that interval is λ

Applications:

- The number of deaths by horse kicking in the Prussian army (first application in reliability engineering (1898)

:)

- birth defects and genetic mutations
- numbers of errors in a computer code / page
- hair found in McDonald's hamburger
- spread of an endangered animal in Africa
- failure of a machine in one month

Exercise: A life insurance salesman sells on average 4 life insurance policies per week. Use Poisson's distribution to calculate the probability that in a given day, he will sell

a) Some policies

(b) 2 or more policies but less than 5 policies

Poisson Distribution as a limit of the binomial

A binomial r.v. with a very small probability of success (p) and very large number of trials n can be approximated with a Poisson distribution with $\lambda = np$, i.e.

$$\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \approx e^{-\lambda} \frac{\lambda^x}{x!}$$

If $n \to \infty$, $p \to 0$, $\lambda = np$

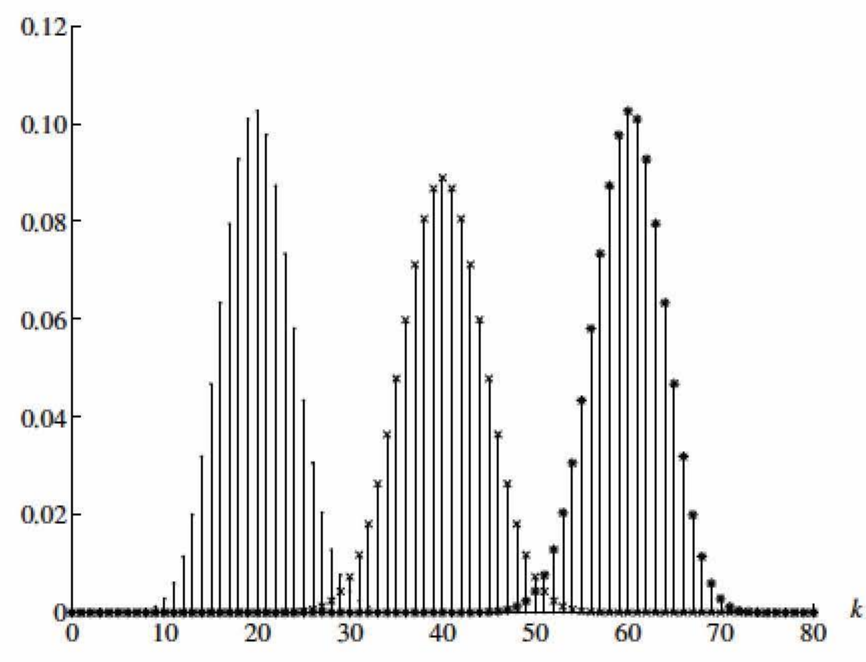
Theorem. Assume that $X_n \sim Bin(n, \frac{\lambda}{n})$ $\forall n$. Also, let $X \sim Pois(\lambda)$.

Then:

$$\lim_{n\to\infty} \mathbb{P}(X_n = x) = \mathbb{P}(X = x) = p_X(x) \quad \forall x \in \{0, 1, 2, \dots\}$$

Proof:

The binomial pmf $p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$ is shown below for $n = 80$ and $p = .25$, $.5$, and 0.75 from left to right, respectively.



Example: Assume that 3% of the tires manufactured in a particular factory are defective. What is the probability of obtaining 5 defective tires from a sample of 50 tires?

Using binomial

$$\mathbb{P}(X=5) = \binom{50}{5} (.03)^{5} (.97)^{45}$$

$$\approx 0.013074$$

Using Poisson $\lambda = np = 50 \times .03 = 1.5$

$$\mathbb{P}(X=5) = \frac{e^{-1.5}(1.5)^5}{5!}$$

$$= 0.014120$$

(f) Negative Binomial

A negative binomial random variable represents the number of independent Bernoulli trials that take place <u>UNTIL</u> success is seen k times.

$X \sim NegB(k, p)$

$$\underbrace{}_{x-1 \text{ trials}}^{k-1 \text{ successes}} \quad \underbrace{}_{x^{th} \text{ trial}}^{k^{th} \text{ success}}$$

Binomial

$P_X(x) =$

Remark: A negative binomial trial can be seen as k (independent) Geo_1 trials (looking for the first success, k times)

Example: What is the probability that we need 100 coin tosses before we see the 90^{th} head?

(g) Power Law Distributions:

Power law distributions are distributions which are proportional to $x^{-\alpha}$

(i) A general power law distribution is the

Zipf_ Mandelbrot Distribution:

$$P_X(x) = \frac{c}{(x+q)^{\alpha}} \qquad x = 1, 2, \dots, N$$

Question: What is the value of c?

(ii) when $N \to \infty$, the pmf becomes the Hurwitz Zeta function $\zeta(\alpha, q)$.

$$P_X(x) = \frac{c}{(x+q)^\alpha} \qquad x = 1, 2, \ldots$$
$$\alpha > 1$$

For this function to be a pmf, α must be bigger than 1. (Why?)

iii) when $N \to \infty$, $q = 0$, we have the zeta - distribution

$X \sim \underbrace{\text{Zeta}}_{\zeta}(\alpha)$

$$P_X(x) = \frac{c}{x^{\alpha}} \qquad x = 1, 2, \ldots$$

$\alpha > 1$

iv) when $q = 0$ and N is finite, we have the Zipf distribution

$$P_X(x) = \frac{c}{x^{\alpha}} \qquad x = 1, 2, \ldots, N$$

($\alpha > 1$ is not needed)

$X \sim \text{Zipf}(\alpha)$

Example: If $X \sim Zeta(2)$, determine the pmf of X

$$P_X(x) = \frac{c}{x^2}$$

$$\sum_{x=1}^{\infty} \frac{c}{x^2} =$$

(v) Another power law

(We use $X \sim Pow(\alpha)$)

$$P_X(x) = \frac{1}{x^{\alpha}} - \frac{1}{(x+1)^{\alpha}} \qquad x \in \mathbb{N}, \quad \alpha > 0$$

Exercise: Show that

$$\sum_{x=1}^{\infty} \left(\frac{1}{x^{\alpha}} - \frac{1}{(x+1)^{\alpha}} \right) = 1 \quad \forall \alpha > 0$$

h) Hypergeometric Distribution

This probability distribution represents a random variable that gives the number of successes (x) in n draws,

without replacement, from a finite population of size N that contains exactly K successes, given that each draw is a success or failure

$X \sim HG(N, K, n)$

$$P_X(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$x \in \mathbb{N}$, for $P_X(x) > 0$, we must have $x \leq \min(K, n)$

$x \geq \max(0, n+K-N)$

To prove that $\sum P_X(x) = 1$, we should use the Vandermond Inequality:

Proposition: $\displaystyle\sum_{x=0}^{n} \binom{K}{x}\binom{N-K}{n-x} = \binom{N}{n}$

Proof.

$$(1+a)^N = (1+a)^K (1+a)^{N-K}$$

Using binomial theorem for the RHS:

When $N \to \infty$, the effect of sampling without replacement becomes less and less. One can imagine that drawing samples from the population

N has a success probability of $\boxed{\quad\quad}$, and it can be modeled as a $\boxed{\quad\quad}$ distribution with parameters $\boxed{\quad\quad}$

Theorem: Assume that

$X_N \sim HG(N, pN, n)$ and

$X \sim Bin(n, p)$. Then

$$\lim_{N \to \infty} \mathbb{P}(X_N = x) = \mathbb{P}(X = x)$$

Proof: Too involved. Will be handed out.

Multinomial Distribution: It is a generalization of the binomial distribution, where instead of two outcomes, each trial has r outcomes with probabilities $p_1, p_2, \dots, p_r$. The probability of seeing x_i of outcome i is:

$$P(x_1, x_2, \dots, x_r) = \frac{n!}{x_1! x_2! \cdots x_r!} p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r}$$

where $x_1 + x_2 + \cdots + x_r = n$ $p_1 + p_2 + \cdots p_r = 1$

Example: Rolling a six-sided die n times.

Dealing with Multiple Random Variables

Joint pmfs:

In reality, we deal with collections of r.v.'s, rather than just a single r.v.

The relationship between two discrete r.v.'s X and Y is captured by their joint pmf.

$$P_{X,Y} : \mathbb{R}^2 \to [0,1]$$

$$P_{X,Y}(x,y) = \mathbb{P}$$

More generally, if $X_1, \ldots, X_n$

are r.v.'s:

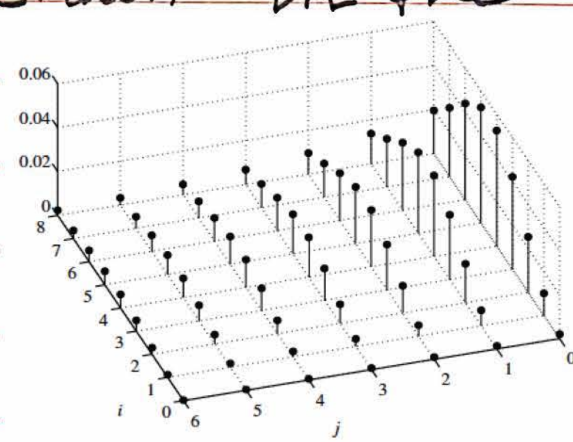
$$P_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \mathbb{P}$$

Example: 100 independent coin tosses. Define X = number of heads and Y = number of tails

$P_{XY}(1,1) = \mathbb{P}$

$P_{X,Y}(1,99) =$

An example of a joint pmf for two integer valued r.v.'s X, Y is shown in the following:



Marginal pmfs:

If X, Y are two r.v.'s with joint pmf $p_{X,Y}$, then their individual pmfs p_X and p_Y are called the marginal pmfs, and

can be retrieved using the following equations:

$p_X(x) =$

$p_Y(y) =$

Reason: The law of Total Probability

Conditional pmfs

Let X and Y be two discrete r.v.'s on $(\Omega, \mathcal{F}, P)$. The <u>conditional</u> pmf of X given Y, $P_{X|Y}$ is defined as:

$p_{X|Y}(x|y) =$

given that $\mathbb{P}(Y=y)$ $\boxed{}$

Remark: If $\mathbb{P}(Y=y)=0$, we leave $P_{X|Y}(x|y)$ $\boxed{}$

Therefore

$$P_{X|Y}(x|y) = \frac{\mathbb{P}}{\mathbb{P}} =$$

where $P_Y(y) \neq 0$ > 0

More generally, if $X_1, X_2, \dots, X_n$ and $Y_1, \dots, Y_m$ are r.v.'s defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we define

$$P_{X_1, \dots, X_n | Y_1, \dots, Y_m}(x_1, x_2, \dots, x_n | y_1, y_2, \dots, y_m)$$

$$= \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | Y_1 = y_1, Y_2 = y_2, \dots, Y_m = y_m)$$

$$= \frac{\quad}{\quad}$$

Example: The joint pmf of X,Y is given:

$$p_{X,Y}(x,y) = \begin{cases} \dfrac{2}{n(n+1)} \left(\dfrac{x}{x+1}\right)^y & x=0,\dots,n-1 \\ & y=0,1,2,\dots \\ 0 & \text{otherwise} \end{cases}$$

Find $p_{Y|X}(y|x)$.

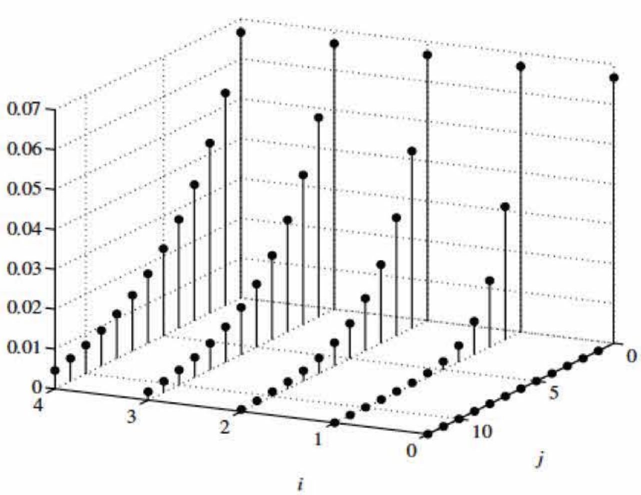


***Figure* 3.3.** Sketch of bivariate probability mass function $p_{XY}(i,j)$ of Example 3.10 with $n=5$. For fixed i, $p_{XY}(i,j)$ as a function of j is proportional to $p_{Y|X}(j|i)$, which is $\text{geometric}_0(i/(i+1))$. The special case $i=0$ results in $p_{Y|X}(j|0) \sim \text{geometric}_0(0)$, which corresponds to a constant random variable that takes the value $j=0$ with probability one.

Exercise: Assume that X is a Discrete Uniform R.V., $dU(1, N+1)$, given that N is a Poisson R.V. with parameter λ. Find the pmf of X.

The Law of Total Probability

Let A be an event and X a discrete r.v in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let

$$B_i = \{X = x_i\} = \{\omega \in \Omega \mid X(\omega) = x_i\}$$

where x_i's are the values that X can take (i.e. $R_X = \bigcup_i \{x_i\}$).

B_i's form a partition of Ω (why?)

The law of Total Probability yields:

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)$$

$$= \sum_i \mathbb{P}(A \mid X = x_i)\, \mathbb{P}(X = x_i)$$

$$= \sum_i \mathbb{P}(A \mid X = x_i)\, P_X(x_i)$$

Obviously, $A = \{Y = y_i\}$ yields the formula for marginal probability mass function of som r.v. Y.

Question: How can one formulate the Baye's Rule using events $\{X=x_i\}$ and $\{Y=y_j\}$?

The Substitution Law

Assume that X, Y are discrete r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$ and $Z = g(X,Y)$ where $g: \mathbb{R} \to \mathbb{R}$. The pmf of Z can be calculated as:

$$\mathbb{P}(Z = z) = \mathbb{P}(g(X,Y) = z)$$

$$= \sum_i \mathbb{P}(g(X,Y) = z \mid X = x_i) \underbrace{\mathbb{P}(X = x_i)}_{P_X(x_i)}$$

- The substitution law states that X can be substituted with x_i in the above expression, i.e.

$$\mathbb{P}(Z=z) = \sum_i \mathbb{P}(g(x_i, Y)=z \mid X=x_i)\,P_X(x_i)$$

Proof: Show the substitution law is correct by using

$$\{\omega \mid g(X(\omega), Y(\omega)) = z\} \cap \{\omega \mid X(\omega) = x_i\}$$

$$= \{\omega \mid g(x_i, Y(\omega)) = z\}$$

and apply it to the definition of conditional probability.

Independence of Discrete Random variables.

Recall that X, Y are called independent iff $\mathbb{P}(X \in B_1, Y \in B_2) = \mathbb{P}(X \in B_1) \times \mathbb{P}(Y \in B_2)$.

We also saw that X, Y are independent iff $F_{X,Y}(x, y) = F_X(x) F_Y(y)$ $\forall x, y \in \mathbb{R}$

Theorem: Assume that X, Y are discrete r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$.

The following statements are equivalent:

(a) X, Y are independent.

(b) $\forall x, y \in \mathbb{R}$, the events $\{\omega \mid X(\omega) = x\}$ and $\{\omega \mid Y(\omega) = y\}$ are independent.

(c) $\forall x, y \in \mathbb{R}$, $P_{X,Y}(x,y) = P_X(x)\, P_Y(y)$

(d) $\forall x, y \in \mathbb{R}$ if $P_Y(y) > 0$, then $P_{X|Y}(x|y) = P_X(x)$

Many of the equivalences are obvious. We show (c) $\Rightarrow$ (a).

Exercise: Generalize (a)–(d) in the previous Theorem for multiple r.v.'s $X_1, X_2, \ldots, X_n$

Example: $X_1, X_2, \ldots, X_n$ are independent $\Leftrightarrow P_{X_1 X_2 \cdots X_n} =$

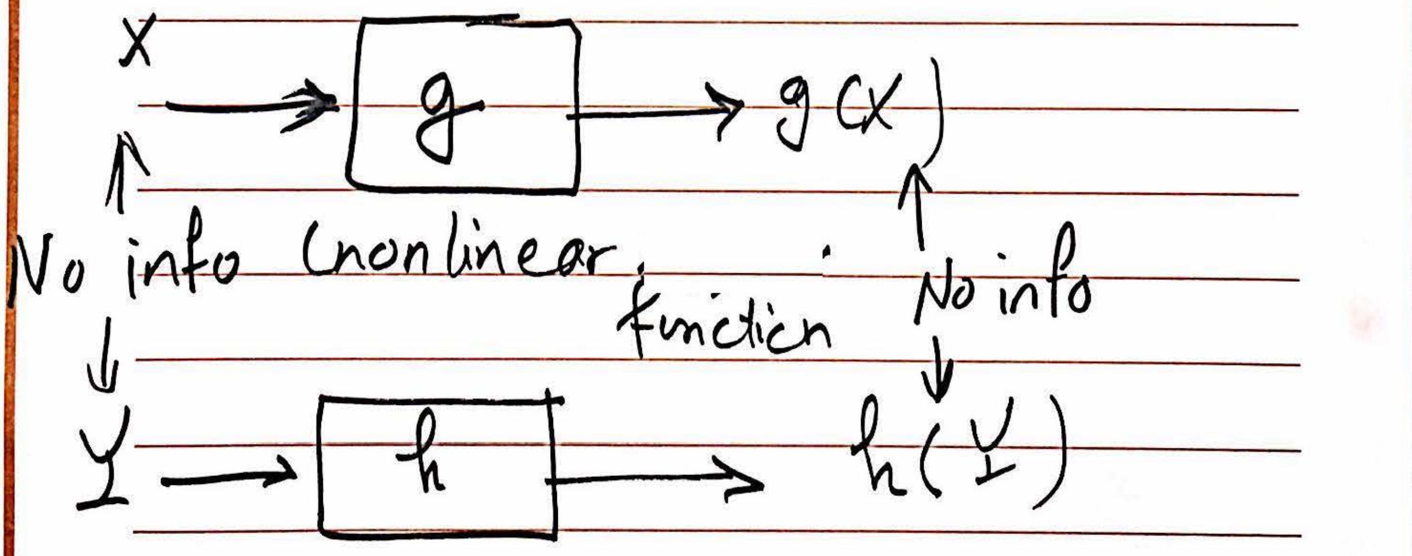
Theorem: Let X, Y be independent discrete r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$.

Let g, h be real functions, i.e. $g: \mathbb{R} \to \mathbb{R}$ and $h: \mathbb{R} \to \mathbb{R}$.

Then, the random variables $g(X)$, $h(Y)$ are independent.

Proof: Exercise.

Interpretation: If X, Y are independent, they contain no information about each other. Passing them through functions g and h does not make the results have any information about each other



Question: Why if X and Y are discrete, $g(X)$ and $h(Y)$ are <u>always</u> random variables even if $g(\cdot)$ and $h(\cdot)$ are not continuous?

Example: n independent coin flips

$X = \#$ of heads

$Y = \#$ of tails $= n - X$

$\mathbb{P}(\{H\}) = p$

Are X, Y independent?

Intuition: X, Y are $\boxed{}$

$\mathbb{P}(X=0) =$

$\mathbb{P}(Y=0) =$

$\mathbb{P}(X=0, Y=0) =$

$\mathbb{P}(X=0, Y=0) \boxed{} \mathbb{P}(X=0)\mathbb{P}(Y=0)$

Derived Distributions
(Functions of Discrete R.V.'s)

Given a discrete random variable with density P_X, and a function $g: \mathbb{R} \to \mathbb{R}$, we are interested in finding the CDF/PMF of the random variable $Y = g(X)$.

$$\mathbb{P}(Y=y) = P_Y(y) = \sum_{\boxed{\{x| \quad \}}} P_X(x)$$

It is a Borel Set because

Example: Assume that $X \sim \text{Pois}(\lambda)$ and $Y = X^2$. Find the pmf of Y.

$$P_Y(y) = \sum_{x \mid x^2 = y} P_X(x)$$

$$= \begin{cases} & y < 0 \\ & y \geq 0, \quad \boxed{} \end{cases}$$

$$= \begin{cases} \\ \\ \end{cases}$$

$$= \begin{cases} \\ \end{cases}$$

More generally, if $X_1, \dots, X_n$ are discrete r.v.'s and $g: \mathbb{R}^n \to \mathbb{R}$ is a function, the pmf of $Y = g(X_1, X_2, \dots, X_n)$ is obtained as

$$P(Y=y) = P_Y(y) = \sum_{\{(x_1, x_2, \dots, x_n) | \quad\quad \}}$$

Sum of Independent Random Variables – Convolution

If X and Y are independent discrete r.v.'s, the pmf of $X+Y$ can be calculated as:

$$Z = X + Y$$

$$P_Z(z) = P_{X+Y}(z) = \mathbb{P}(X+Y=z)$$

$$= \sum_{\{(x,y)\,|\,x+y=z\}} \mathbb{P}(X=x, Y=y) =$$

(Convolution Sum)

Remark: The Convolution Sum can also be derived by the substitution law. Prove it using the substitution law!

Exercise: Assume that

$X \sim Pois(\lambda)$ and $Y \sim Pois(\mu)$.

Show that $Z = X + Y$ has

a Poisson distribution with $\mu + \lambda$,

ie $Z \sim Pois(\mu + \lambda)$

Exercise: Assume that X, Y are discrete r.v./s on $(\Omega, \mathcal{F}, \mathbb{P})$, and are independent. Use the substitution law to find an expression for the pmf of $Z = XY$.

Max and Min Problems

Let $X_1, \dots, X_n$ be independent Random Variables.

The cdf of cdf of

$Y = \max(X_1, X_2, \dots, X_n)$ is

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\max(X_1, \dots, X_n) \leq y)$$

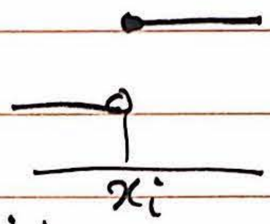
$=$

The CDF of $Z = \min(X_1, X_2, \dots, X_n)$ can be calculated as:

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(\min(X_1, \dots, X_n) \leq z)$$

Remark: The above argument for calculating the CDF of Min and Max of independent r.v.'s is <u>NOT</u> restricted to discrete r.v.'s. In the

Case of discrete r.v.'s, the pmf can easily be calculated from the cdf at its points of discontinuity,



$$\mathbb{P}(X = x_i) = F_X(x_i) - \lim_{y \to x_i^-} F_X(x_i)$$