


Specifically, what is the largest <sup>possible</sup> value  
of  $P(X \geq a)$  while keeping  $E(X) = m$   
the same  
say  $a = 3$  mean value say  $m = 5$

Ans:

pf:

② Inequality = A refinement of the Markov Inequality, which gives you more accurate estimate at the price of requiring more info. Need:



Ex: I am recording the values of my stock portfolio, which has 100 stocks.

One day my computer crashes, I only remember the average price is

50.

Q: At most how many stocks can have the price  $> 90$ ?

Ans:

Q: Suppose I also remember the standard deviation is 20.

What is the maximum number of stocks having value  $\geq 90$ .

Ans:

Q: Suppose I also know it is a bell-shaped distribution  
What is  $P(X > 90)$ ?

Ans:

④ Inequality = A further refinement  
of the Markov inequality  
We need only

---

Chernoff Inequality.

\* Continue from our example. For a  
Gsr w.  $\mu=50$   $\sigma=20$ ,  $a=90$   
$$X^*(s) = e^{-s \cdot \mu + \frac{\sigma^2 s^2}{2}}$$

Pf of the Chernoff bound.

\* Chernoff bound is tricky.

Since once knowing  $X^*(s)$ , we already know the exact distribution of  $X$ . We can thus use summation/integration to find the "exact" value of  $P(X \geq a)$ . However, finding the exact value of  $P(X \geq a)$  is computationally intensive. In many cases, find the Chernoff bound value is just as good.

## Chapter 5 Pairs of R.Vs.

Consider a R.V  $X$  with sample space  $S_X = \{0, 1\}$  and another R.V  $Y$  with  $S_Y = \{0, 1, 2\}$ .

To discuss the joint relationship of  $X$  and  $Y$ , we need to consider the

Once the W.A of  $S_{X,Y}$  is made, we can compute any prob like

$$P(X^2 \leq \sqrt{Y}), P(\max(X, Y) \leq 3) \dots$$

Note that Even  $E(X^2 Y)$

it is no different than considering

a Random Vector  $W = (X, Y)$  where

the output of each  $W$  is a vector.

Nonetheless, it is not efficient to always assign the Weight for the joint sample space as most of the time, we are interested only in  $P(X > 0)$ ,  $P(\sqrt{X} < 3)$ ,  $E(X^{\frac{3}{2}})$ . In most cases, we do the W.A gradually, start from the W.A of  $X$ , then extend to  $(X, Y)$ . We say the W.A for 'X' alone is

the

	Y	0	1	2
X	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
	1	$\frac{1}{4}$	0	$\frac{1}{4}$

joint distribution

the key concept of every imputation

Note 1: marginal + marginal  $\Rightarrow$  joint. 138

		Y								
		0	1	2						
Ex	x	<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"><tr><td style="padding: 5px;">5/24</td><td style="padding: 5px;">2/24</td><td style="padding: 5px;">5/24</td></tr><tr><td style="padding: 5px;">5/24</td><td style="padding: 5px;">2/24</td><td style="padding: 5px;">5/24</td></tr></table>			5/24	2/24	5/24	5/24	2/24	5/24
	5/24				2/24	5/24				
5/24	2/24	5/24								
	0									
	1									

has the same marginal distributions as the first example. but different joint distribution.

---

Ex HW10Q8

X is geometric with p.

↳ The marginal distribution

Given  $X=x_0$ , Y is a Poisson with

$$\lambda = x_0$$

↳ The conditional distribution

Q: The joint sample space = ?

Ans:

Q: What is the joint W.A ?