④ Chernoff Inequality: A further refinement of the Markov inequality

We need only ④ The moment generating function $X^*(s)$

Chernoff Inequality.

For any a

$$p(x \geq a) \leq e^{sa} X^*(s)$$

for any negative $s \leq 0$

To have the tightest bound

$$\leq \min_{s \leq 0} e^{sa} X^*(s)$$

Take the min over all possible $s \leq 0$.

* Continue from our example. ~~For~~ a Gsn w. $\mu = 50$ $\sigma = 20$, $a = 90$

$$X^*(s) = e^{-s \cdot \mu + \frac{\sigma^2 s^2}{2}}$$

$$= e^{-50s + 200s^2}$$

$$e^{sa} \cdot X^*(s) = e^{90s - 50s + 200s^2} = e^{40s + 200s^2}$$

$$= e^{200(s+\frac{1}{10})^2 - 2}$$

$\Rightarrow$ The min value is when $s = -\frac{1}{10}$

$$\Rightarrow P(X \geq 90) < e^{-2} \approx 0.135$$

Pf of the Chernoff bound. for any $s \leq 0$

$$P(X \geq a) = P(-sX \geq -sa)$$

$$= P(e^{-sX} \geq e^{-sa}) \leq \frac{E(e^{-sX})}{e^{-sa}} = e^{sa} X^*(s)$$

$*$ 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 $\boxed{\text{joint sample space}}$

$$S_{XY} = \{ (0,0)^{\frac{1}{6}}, (0,1)^{\frac{1}{6}}, (0,2)^{\frac{1}{6}}, (1,0)^{\frac{1}{4}}, (1,1)^{0}, (1,2)^{\frac{1}{4}} \}.$$

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

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

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 $\boxed{\text{marginal distribution.}}$

X \ Y	0	1	2	
0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
	$\frac{5}{12}$	$\frac{2}{12}$	$\frac{5}{12}$	

marginal distribution of X

marginal distribution

Marginal distribution

$\Uparrow$

$\boxed{\text{joint distribution}}$

$\Downarrow$

Conditional distribution

marginal + conditional

$\Downarrow$

joint

the key concept of every computation

Note 1: marginal + marginal $\not\Rightarrow$ joint.

Ex

$$\begin{array}{c|ccc|c} X \backslash Y & 0 & 1 & 2 & \\ \hline 0 & 5/24 & 2/24 & 5/24 & \frac{1}{2} \\ 1 & 5/24 & 2/24 & 5/24 & \frac{1}{2} \\ \hline & 5/12 & 2/12 & 5/12 & \end{array}$$

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

Ex HW10Q8

X is geometric with p.

$\longrightarrow$ The marginal distribution

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

$\longrightarrow$ The conditional distribution

Q: The joint sample space = ?

Ans: $S_{XY} = \{$ all pairs of non-negative integers

$(0,0), (0,1), (0,2) \cdots$

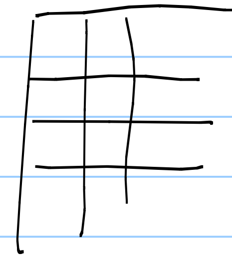
$(1,0), (1,1) \cdots$

$\vdots \quad \ddots$

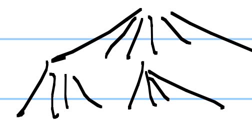
Q: What is the joint W.A ?

$$P(X=k, Y=h)$$

$$= \underbrace{P(Y=h \mid X=k)}_{\text{joint}} \cdot \underbrace{P(X=k)}_{\text{marginal}}$$



(think about the tree method)
table



$$= \begin{cases} \left(\dfrac{k^h}{h!} e^{-k}\right) \cdot p(1-p)^k & \text{if } 0 \le h, \ 0 \le k \\ 0 & \text{otherwise} \end{cases}$$

Q: What is the marginal W.A of X ?

A: Summing over all Y values

$$P(X=k) = \sum_{h=0}^{\infty} P(X=k, Y=h)$$

$$= \begin{cases} p(1-p)^k \underbrace{\sum_{h=0}^{\infty} \dfrac{k^h}{h!} e^{-k}}_{=1} & \text{if } k \ge 0 \\ 0 & \text{if } k < 0. \end{cases}$$

Q: How to compute the marginal distribution of Y ?

Ans: Summing over X's

$$P(Y=h) = \sum_{k=0}^{\infty} P(X=k, Y=h)$$

$$= \begin{cases} \displaystyle\sum_{k=0}^{\infty} \frac{k^h}{h!} e^{-k} p(1-p)^k \\ \quad = \displaystyle\frac{p}{h!} \sum_{k=0}^{\infty} k^h e^{-k} (1-p)^k & \text{if } h \geq 0 \\ 0 & \text{if } h < 0 \end{cases}$$

The computation of the sum is not easy.

Ex: $P(Y=0) = \frac{p}{0!} \sum_{k=0}^{\infty} (k)^0 e^{-k} (1-p)^k$

$$= p \times \frac{1}{1 - e^{-1}(1-p)}$$

$$P(Y=1) = \frac{p}{1!} \sum_{k=0}^{\infty} k e^{-k} (1-p)^k$$

$$= \frac{p}{1!} \times \frac{e^{-1}(1-p)}{(1-e^{-1}(1-p))^2}$$

$\boxed{\text{* The geometric series formulas}}$

* But the concept is straightforward

Q: $P(X^2 + Y^2 \leq 4)$

Ans: $= P\left(\begin{matrix} (0,0), & (0,1), & (0,2) \\ (1,0) & (1,1) & \\ (2,0) & & \end{matrix} \right) = 6$ terms.

Independence

Two (marginal) R.Vs X and Y are independent. If their joint weight assignment is the product of the marginal probabilities.

That is

$$P(X=k, Y=h) = P(X=k)\cdot \boxed{P(Y=h)}$$

vs.

Comparison, if not independent

$$p(X=k, Y=h) = P(X=k)\cdot \boxed{P(Y=h \mid X=k)}$$

$\Rightarrow$ marginal + conditional $\Longrightarrow$ joint

marginal + marginal + indep $\Longrightarrow$ joint

Note: Relate it to the tree/table method we have learned so far.

* Discrete 2-dim R.V.s.

$$S_{XY} = \left\{ \begin{matrix} (-1,0) & (-1,1) & \cdots \\ (0,0), & (0,1) & \\ (1,0), & \cdots & \end{matrix} \right.$$

all grid points $\}$.

The Joint W.A is specified by the joint pmf: $p_{k,h} = P(X=k, Y=h)$

s.t. $p_{k,h} \geq 0$

$$\sum_{k=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} p_{k,h} = 1.$$

* From joint distribution to marginal distri

$$P(Y=h) = \sum_{k=-\infty}^{\infty} p_{k,h}$$ summing over uninterested variable

focusing on the column

* Expectation of a function of a 2-dim R.V

$E(f(X,Y))$ say $f(x,y) = x^2 + y^2 + xy$

$$\text{is } \sum_k \sum_h f(k,h) \underbrace{p_{k,h}}_{\text{weight}}$$

$E(X^2 + Y^2 + XY)$

Q: $E(U(1-X-Y))$

$$U(1-x-y) = \begin{cases} 1 & \text{if } 1-x-y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Ans $= 1 \times p_{0,0} + 1 \times p_{0,1} + 1 \times p_{10} + 0 \cdots$

$= 3$ terms

HW10Q8

X is geometric with p.

$\longrightarrow$ The marginal distribution

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

$\alpha = x_0$

$\longrightarrow$ The conditional distribution

Q: $E(X) = ?$

Ans: $\frac{1-p}{p}$ by ~~table~~ look-up since the (marginal) distribution of X is geometric

Q: $E(Y) = ?$

Ans: We can compute either

① $\sum_{y=0}^{\infty} y P(Y=y)$ hard, we don't even know $P(Y=y)$

Or ② $\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} y P(X=x, Y=y) = \sum_{y=0}^{\infty} y \sum_{x=0}^{\infty} P(X=x, Y=y)$

The difference is how you compute the weighted average, in a "① Column by column" way or "② block by block" way

from ② $\Rightarrow \sum_{x=0}^{\infty} \underbrace{\sum_{y=0}^{\infty} y \cdot \frac{x^y}{y!} e^{-x}}\, p(1-p)^x$

$$= \sum_{x=0}^{\infty} \underbrace{\underbrace{x}_{\text{by table look up for the expected value of a Poisson}} \quad p(1-p)^{x}}$$

$$= \frac{1-p}{p} \quad \text{by another table look up.}$$

Summary: For 2-dim R.Vs, there are many ways of "counting the weights". Some are easier than the others.

* Conditional distribution for discrete variables.

$P(X=k \mid Y=h)$ focusing on the h-th col.

$$= \frac{P(X=k, Y=h)}{P(Y=h)}$$

$$= \frac{p_{k,h}}{\sum_{x=-\infty}^{\infty} p_{x,h}}$$

Exercise: $P(Y=h \mid X=k) = \dfrac{p_{k,h}}{\sum_{y=-\infty}^{\infty} p_{k,y}}$

Note Title 3/9/2011

Summary

* Discrete 2-dim R.Vs.

$S_{XY} = \{\text{all grid points}\}$

W.A: $p_{k,h} = P(X=k, Y=h)$ (2-dim) joint pmf

Expectation $E(g(X,Y))$

$$= \sum_k \sum_h g(k,h)\, p_{k,h}$$

Marginal pmf $P(X=k) = p_k = \sum_{h=-\infty}^{\infty} p_{k,h}$

summing over the row of interest.

$$P(Y=h) = p_h = \sum_{k=-\infty}^{\infty} p_{k,h}$$

Conditional pmf $p_{k|Y=h} = P(X=k \mid Y=h)$

$$\triangleq \frac{p_{k,h}}{\sum_{x=-\infty}^{\infty} p_{x,h}}$$

Conditional expectation ?

Conditional pmf + expectation.

Note Title 3/9/2011

* Continuous 2-dim Random variables.

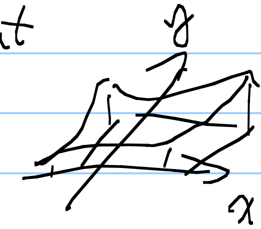
$S_{XY} = \{$ all pairs of real numbers

$(-\infty, \infty) \times (-\infty, \infty)$

say $(0.1, 1001.5), (\pi, e) \ldots$

The joint W.A is specified by the joint

pdf: $f_{XY}(x,y)$ such that



$f_{XY}(x,y) \geq 0$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1$$

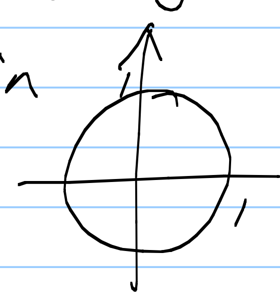
$P((X,Y) \in A)$

$\in$ ↳ falls in / belongs to

$$= \int_A f_{XY}(x,y)\,dx\,dy$$

Problem 5.28(i)

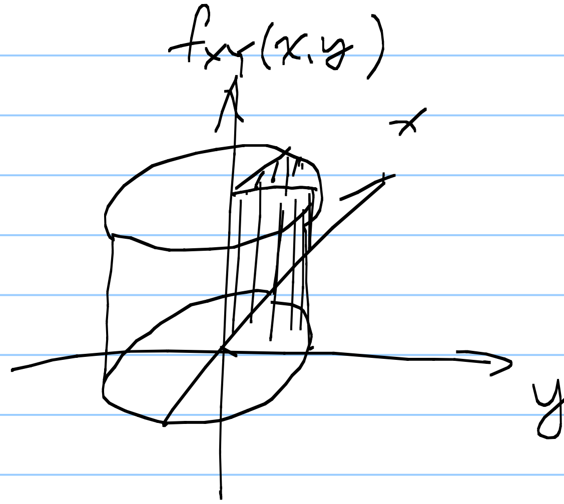
Ex: HW11Q3. A joint pdf is given as follows $f_{XY}(x,y) = \begin{cases} k & \text{if } (x,y) \text{ in} \\ 0 & \text{otherwise.} \end{cases}$



Q: find the k value.

$\Rightarrow k \times \pi \times 1^2 = 1 \Rightarrow k = \frac{1}{\pi}.$

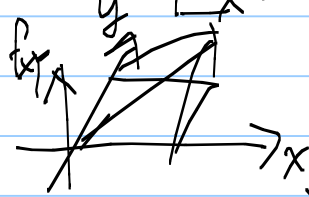
Q: Find $P(X>0, Y>0)$?



Ans: $= k \times \left(\frac{1}{4}\pi \times 1^2\right) = \frac{1}{4}$

---

* Another example.



Ex: $f_{XY}(x,y) = \begin{cases} kxy & \text{if } 0<x<1 \\ & \;\&\; 0<y<1 \\ 0 & \text{otherwise} \end{cases}$

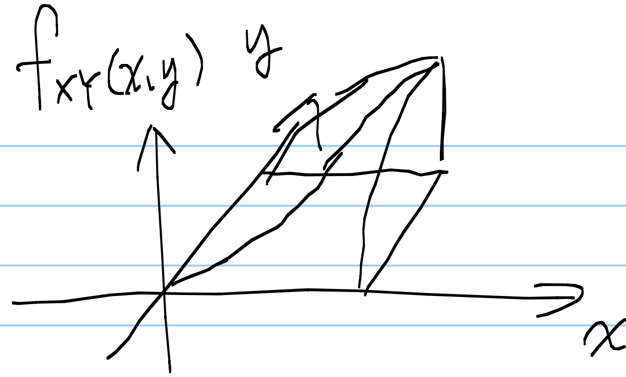
is a joint pdf, ① find k value.

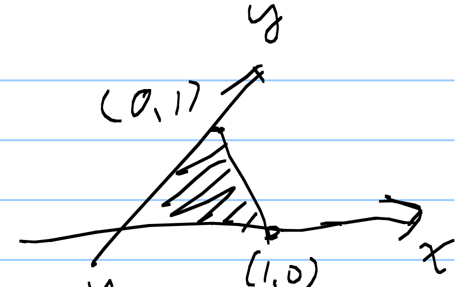
② $P(X+Y \leq 1)$

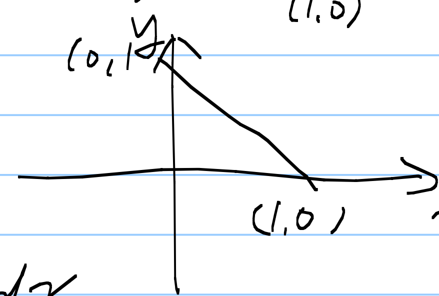
Ans: ① $\int_0^1 \int_0^1 kxy\,dx\,dy = 1$

$\Rightarrow \int_0^1 ky\left(\frac{x^2}{2}\Big|_0^1\right) dy$ $\Rightarrow \boxed{k=4}$

$= \int_0^1 \frac{1}{2} ky\,dy = \frac{1}{2}k\left(\frac{y^2}{2}\Big|_0^1\right) = \frac{k}{4}$







Q2:

$$P(X+Y\leq 1)=\int_{x=0}^{1}\int_{y=0}^{1-x} 4xy\,dy\,dx$$

$$=\int_{x=0}^{1} 4x\left(\frac{y^2}{2}\Big|_0^{1-x}\right)dx$$

$$=\int_{x=0}^{1} 4x\cdot\frac{(1-x)^2}{2}\,dx$$

$$=\int_{x=0}^{1} 2x-4x^2+2x^3\,dx$$

$$=x^2-\frac{4}{3}x^3+\frac{2}{4}x^4\Big|_0^1=\frac{1}{6} \quad \#$$

Exercise $P(X+Y\leq 1.5)=?$

Ans: $\iint_{\square} 4xy\,dy\,dx$