第 二 章

随机变量及其分布

(Part 2)

第四节：数学期望

在许多实际问题中：

(1) 不需要知道 X 的精确概率分布，只需要知道它的某些特征；

(2) 可能难于求出 X 的精确概率分布，退而求其次。

希望知道随机变量的下列特征：

(1) 集中的位置；

(2) 集中的程度；

(3) 多个 RV 之间的关系，关系的强弱；

(4) …

一、数学期望的概念

例:甲、乙两人，赌技相同，各出赌注50元，无平局。先赢3局者获得全部赌本100元。

现假设甲赢2局，乙赢1局，终止赌博。

问如何合理分配赌本?

一、数学期望的概念

例: 甲、乙两人，赌技相同，各出赌注50元，无平局。先赢3局者获得全部赌本100元。现假设甲赢2局，乙赢1局，终止赌博。问如何合理分配赌本？

直观1： 基于历史 --- 按已赌局数的胜负局数分，则甲分总赌本的2/3、乙分总赌本的1/3。

合理吗？

一、数学期望的概念

设想：再赌下去，至多再赌两局必可结束（基于未来）

甲，乙甲，乙乙 （红色表示赢）

此三种结果不等可能。为此，可虚拟（第二局）

甲(甲)，甲(乙)，乙甲，乙乙

设 X 为甲所得赌资，则 X 为RV，具有分布

X	0	100
P	¼	¾

甲的基于未来的“期望”（平均）所得为：

$0* \frac{1}{4} + 100* \frac{3}{4} = 75$（元）.

例（投资问题）：设有两股票

问：买哪种股票更理想？（不考虑风险）

例（投资问题）：设有两股票

问：买哪种股票更理想？（不考虑风险）

比较A，B在未来的“平均”价格：

A：110*0.7 + 90*0.3 = 77 + 27 = 104;

B：120*0.7 + 80*0.3 = 84 + 24 = 108;

不考虑风险，应买 B.

注：这里的“平均”不但由 RV 的取值决定，

而且由取该值的概率决定，即是“加权平均”。

二、定义

（离散型RV） 设离散型RV X 的分布列为：

$$P(X=x_k)=p_k,\ k=1,2,\ldots$$

如果级数 $\sum_{k=1}^{+\infty} x_k p_k$ 绝对收敛，即 $\sum_{k=1}^{+\infty} |x_k| p_k < +\infty$,

则称 $E(X) := \sum_{k=1}^{+\infty} x_k p_k$ 为X的数学期望(或均值)。

加权平均

若级数 $\sum_{k=1}^{+\infty} |x_k| p_k$ 不收敛(即使 $\sum_{k=1}^{+\infty} x_k p_k$ 本身收敛),

则称X的数学期望不存在。

注记：

（1）定义中要求级数 $\sum x_k p_k$ **绝对收敛**”是为保证其“和数”**不随级数项的排列次序而发生改变**，保证期望的唯一性。

（如果只是条件收敛，会发生什么？）

→ The value of the sum depends on the order in which x_j are listed.

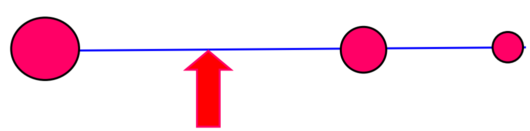
（2）“均值”相当于物理中的“重心”(center of gravity)：

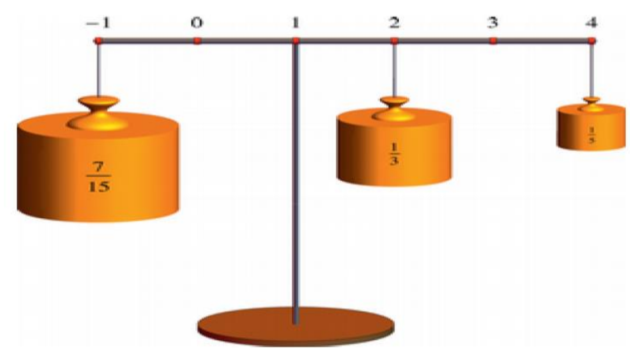
若在x_k处放置质量p_k，则其“重心”位置m满足

$$\sum (x_k - m)\cdot p_k = 0\ (m\text{右边的力矩等于它左边的力矩})$$

$$\Rightarrow m = \sum x_k \cdot p_k.$$

能保持平衡



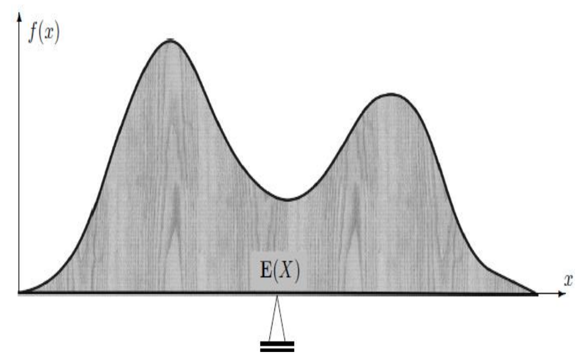


二、定义（连续型RV的数学期望）

设$X \sim f(x)$，考虑用离散型RV近似X：

在$(x, x+\Delta x]$上取值x，概率$f(x)\Delta x$.

该离散型RV的数学期望为：$\sum x\ f(x)\Delta x$.



设连续型RV X的概率密度函数为$f(x)$.

若积分$\int_{-\infty}^{+\infty} x\ f(x)dx$绝对收敛，即 $\int_{-\infty}^{+\infty} |x|\ f(x)dx < +\infty$,

则称 X的数学期望存在，且$E(X) := \int_{-\infty}^{+\infty} x\ f(x)dx$.

若积分$\int_{-\infty}^{+\infty} |x|\ f(x)dx$不收敛（即使$\int_{-\infty}^{+\infty} x\ f(x)dx$收敛），

则称X的数学期望不存在。

例：运气轮

赌徒押注于1到6之间某个数, 然后庄家掷3枚骰子. 如果赌徒押注的这个数出现k次(k=1, 2, 3)，则他赢得k元；如果赌徒押注的数没有出现，则他损失1元。问这个赌博是否公平?

解：设赌徒押的数出现的次数为X,则$X \sim B(3,1/6)$.

Y为赌徒赢得的数目。

则 $P(Y=-1)=P(X=0)=C_3^0(1/6)^0(5/6)^3=125/216.$

$P(Y=1)=P(X=1)=C_3^1(1/6)^1(5/6)^2=75/216.$

$P(Y=2)=P(X=2)=C_3^2(1/6)^2(5/6)^1=15/216.$

$P(Y=3)=P(X=3)=C_3^3(1/6)^3(5/6)^0=1/216.$

$$E(Y)=(-1)\cdot\frac{125}{216}+1\cdot\frac{75}{216}+2\cdot\frac{15}{216}+3\cdot\frac{1}{216}=-\frac{17}{216}.$$

长期赌下去，每216局，赌徒将要输掉17元.

例. 核酸检测新方法

N个人核酸检测，检验是阳性的概率都是 p，独立。

方法一： N 个人的样本分别检验，共需 N 次检验；

方法二： 分组检验，K 个人的样本在一起检验，
若是阴性，则全组均为阴性；
若是阳性，则再分别检验。

比较两种方法的效率（检验次数）。

例. 核酸检测新方法

对方法二（混检），每个人的检验次数为随机变量X,

X的分布列为(设$q=1-p$)　　q 为阴性概率

X	$\frac{1}{k}$	$1+\frac{1}{k}$
P	q^k	$1-q^k$

X的数学期望为：

$$EX=\frac{1}{k}\cdot q^k+(1+\frac{1}{k})\cdot(1-q^k)=1-(q^k-\frac{1}{k}).$$

当$q^k-\frac{1}{k}>0$时，即可减少验血次数。

当p已知时，可选取k,最小化EX。

例如：
p=0.01,k=10,
EX =0.196

例. 核酸检测新方法


```r
p <- c(0.001, 0.01, 0.1)
k <- 1:40

q <- 1 - p[1]
Ex1 <- 1 - q^k + 1/k
which.min(Ex1)

## [1] 32

q <- 1 - p[2]
Ex2 <- 1 - q^k + 1/k
which.min(Ex2)

## [1] 11

q <- 1 - p[3]
Ex3 <- 1 - q^k + 1/k
which.min(Ex3)

## [1] 4
```

三、常见离散型RV的数学期望

1. 两点分布

X	0	1
P	1-p	p

则 E(x) = p.

2. 二项分布

设$X \sim B(n,p)$,即$P(X=k)=C_n^k p^k(1-p)^{n-k}, k=0,1,2,\ldots,n.$

$$\text{则} E(X)=\sum_{k=0}^{n} k C_n^k p^k (1-p)^{n-k}$$

$$=\sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$=np\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}$$

$$=np[p+(1-p)]^{n-1}=np.$$

3. 几何分布

设$X \sim Ge(p)$,即$P(X=k)=p(1-p)^{k-1}, k=1,2,\ldots$

则$E(X)=\sum_{k=1}^{\infty} kP(X=k)=\sum_{k=1}^{\infty} k\,p(1-p)^{k-1}=p\sum_{k=1}^{\infty} k(1-p)^{k-1}$.

设$S=\sum_{k=1}^{\infty} kt^{k-1}=1+2t+3t^2+\cdots,\ (t=1-p)$

则 $\qquad tS= \qquad t+2t^2+\cdots$

相减:$(1-t)S=1+t+t^2+\cdots=\dfrac{1}{1-t},\quad \Rightarrow S=\dfrac{1}{(1-t)^2}=\dfrac{1}{p^2}$.

$$E(X)=p\sum_{k=1}^{\infty} k(1-p)^{k-1}=p\frac{1}{p^2}=\frac{1}{p}.$$

Or: $\sum_{k=1}^{\infty} kt^{k-1}=\sum_{k=0}^{\infty}\frac{d}{dt}(t^k)=\frac{d}{dt}\left(\sum_{k=0}^{\infty} t^k\right)=\frac{d}{dt}\left(\frac{1}{1-t}\right)=\frac{1}{(1-t)^2}$

4. Poisson 分布

设 $X \sim P(\lambda)$, 即 $P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\ldots$

$$\text{则 } E(X)=\sum_{k=0}^{\infty} k\,P(X=k)=\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}e^{\lambda}=\lambda.$$

参数

5. 期望不存在？

设 $RV\ \ X$取值 $x_k=(-1)^k\cdot\dfrac{2^k}{k}$的概率为 $p_k=\dfrac{1}{2^k},k=1,2,\ldots$

则 $p_k\geq 0$ 且 $\sum\limits_{k=1}^{\infty}p_k=1$,因此确为概率分布列。

虽然 $\sum\limits_{k=1}^{\infty}x_k p_k=\sum\limits_{k=1}^{\infty}(-1)^k\cdot\dfrac{2^k}{k}\cdot\dfrac{1}{2^k}=\sum\limits_{k=1}^{\infty}(-1)^k\cdot\dfrac{1}{k}=-\ln 2\cdot$

但是 $\sum\limits_{k=1}^{\infty}|x_k|p_k=\sum\limits_{k=1}^{\infty}\dfrac{2^k}{k}\cdot\dfrac{1}{2^k}=\sum\limits_{k=1}^{\infty}\dfrac{1}{k}=+\infty.$ → 级数条件收敛

所以数学期望不存在。

6. St. Petersburg Paradox*

- A casino offers a game in which a fair coin is tossed. The pot starts at 2 dollar and is doubled every time a Tail appears.
- The first time a Head appears, the game ends and the player wins whatever is in the pot.
- The player wins

--- 2 dollar if a Head appears on the first toss,

--- 4 dollars if a Tail appears on the first toss, and a Head on the second, …

--- 2^N dollars, where *N-1* Tails are tossed before the first Head appears.

- **What would be fair price to pay the casino for entering the game?**

6. St. Petersburg Paradox*

- **Method I:**

  Let X be the money you win, then

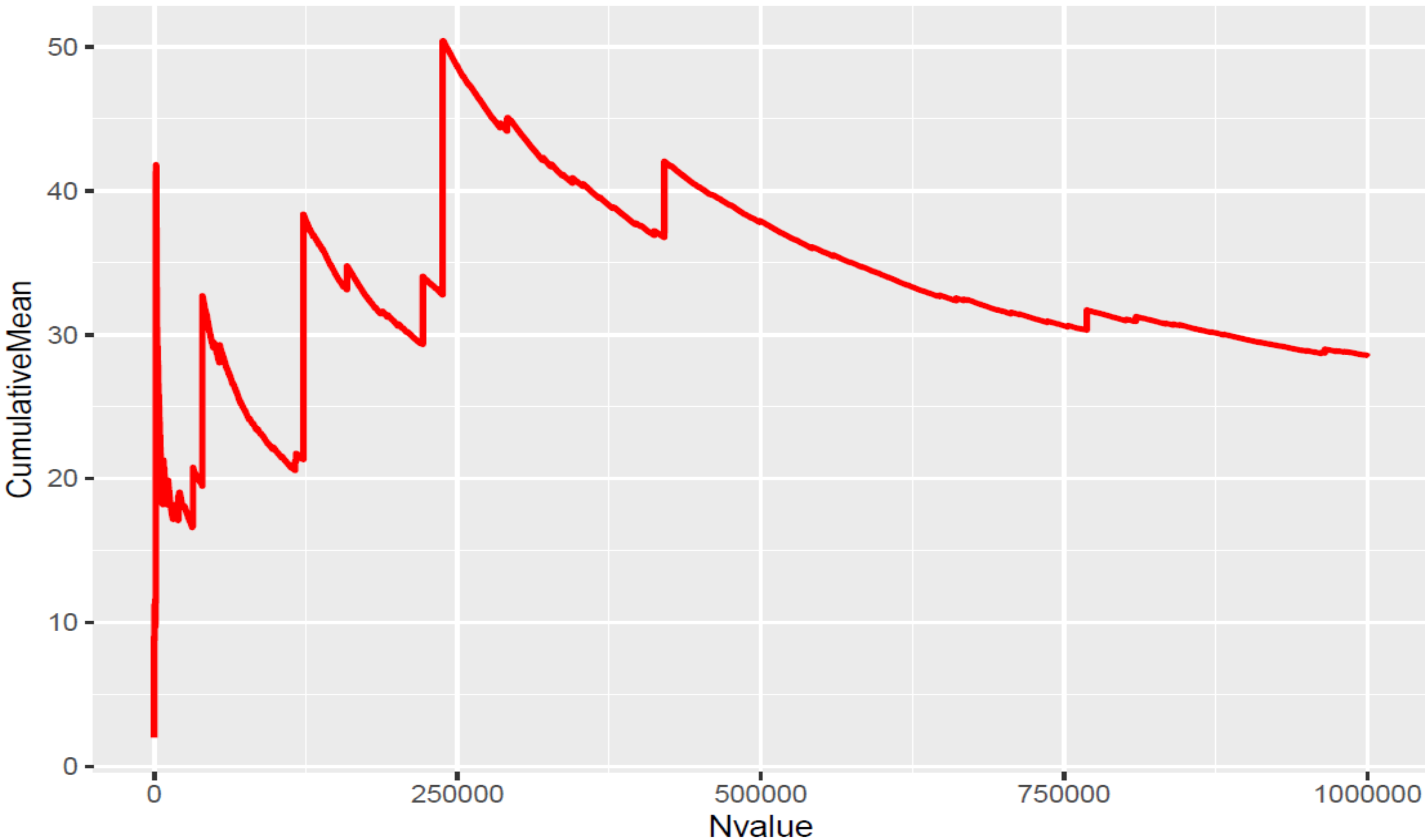
  $E(X) = ½ *2 + ¼*4 + 1/8 *8 + … =$ **infinity**.

  The mathematical expectation does not exist!

- **Method II:**
- Let N be the number of rounds that the game lasts. It is the number of tosses until the first Head, so N ~ Ge(1/2) and E(N) = 2.
- Let X be your winnings from playing the game. By definition, X = 2^N. So

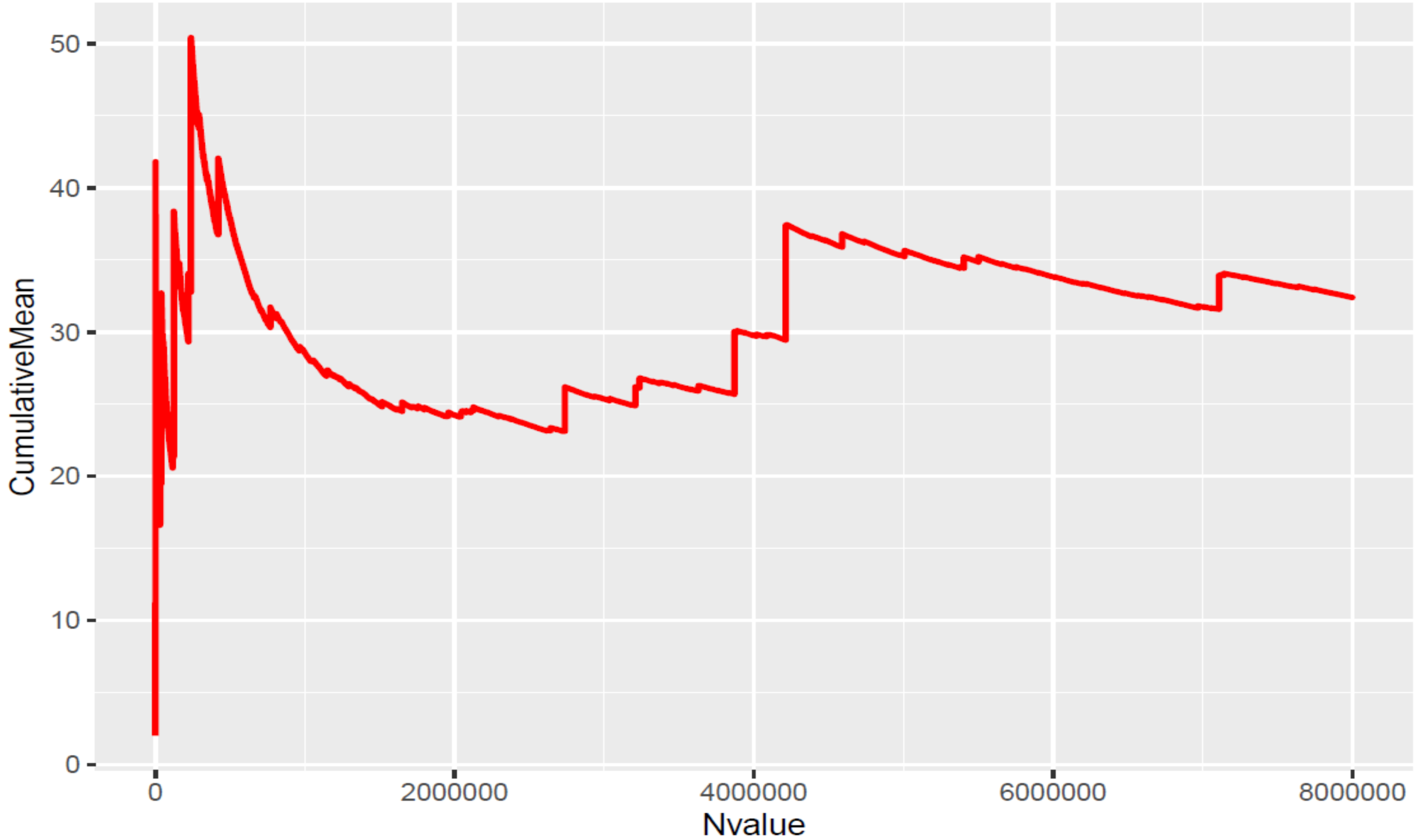
  E(X) = E(2^N) = 2^(E(N)) = 4.   (???)

6. St. Petersburg Paradox*---Simulation



Every once in a while there is a sharp upward jump

6. St. Petersburg Paradox*---Simulation



Every once in a while there is a sharp upward jump

6. St. Petersburg Paradox* --- Expectation?

- Let X be the money you win, then

  E(X) = **infinity**.

  The mathematical expectation does not exist!

- **Paradox!**

Although the expected payoff is infinite, most people would not be willing to pay very much to play the game (even if they could afford to lose the money).

The paradox is that there is **no good answer** directly based on the expected payoff.

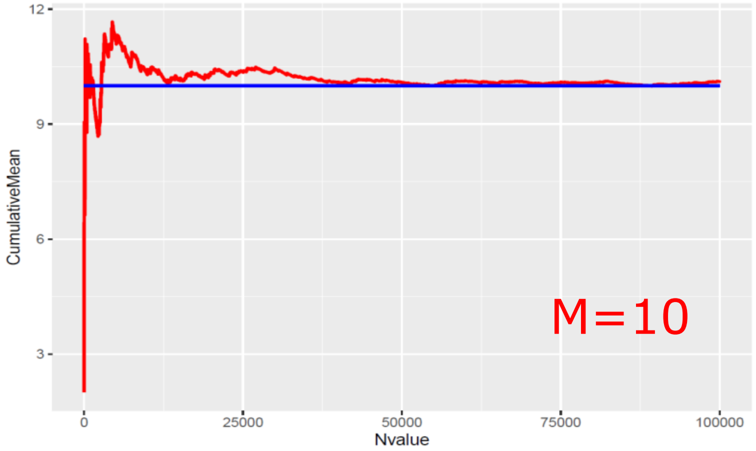
- **What is the problem?**
- **Possible solution?**

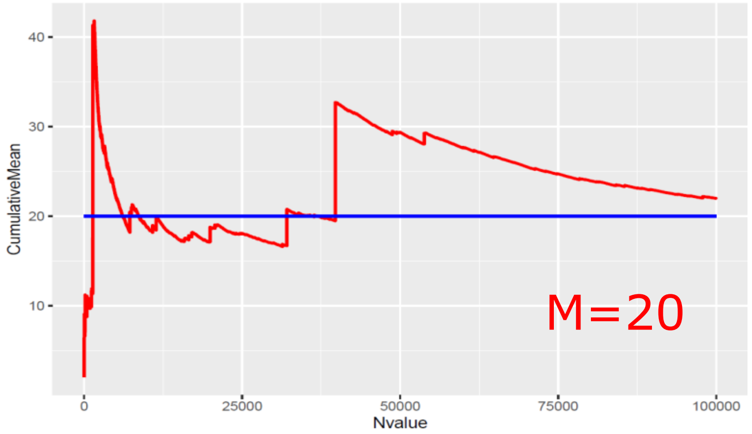
Possible modification to the game

- Given some integer M (say, M=10 or 20). The casino pays player 2^k dollars when heads appears for the first time in the kth toss ($k \le M$), and pays nothing if tails is tossed M times in a row.

$$E(X) = ½ *2 + ¼*4 + \ldots + \frac{1}{2^M} \times 2^M = M.$$

- Simulation: average payoff





Possible solutions * --- Utility Function

- People do not play games as if they are maximizing expected value they receive.
- Rationality assumptions:

  People act as though they are maximizing *something*, called **utility function**.

- **Principle of Decreasing Marginal Utility:**
- An additional unit of money is determined by how much money they already have, or each additional good consumed is less satisfying than the previous one.

  (the value a homeless person would attach to a hundred-dollar bill is far more than Bill Gates would).

四、常用连续型RV的期望

1、均匀分布

设$X \sim U(a,b)$，即概率密度函数为

$$f(x)=\frac{1}{b-a}I_{\{a<x<b\}}=\begin{cases}\dfrac{1}{b-a}, a<x<b\\ 0, \quad \text{其它}\end{cases}$$

则 $E(X)=\int_a^b x\frac{1}{b-a}dx=\frac{a+b}{2}$ ---区间的中点

与直观吻合

2. 指数分布

设$X \sim \mathbf{Exp}(\lambda)$，即$X$的概率密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x \geq 0,\\ 0, & 其它,\end{cases}$$

则 $$E(X)=\int_0^{+\infty} x\lambda e^{-\lambda x}dx=\int_0^{+\infty} x\,d(-e^{-\lambda x})$$

integration by part

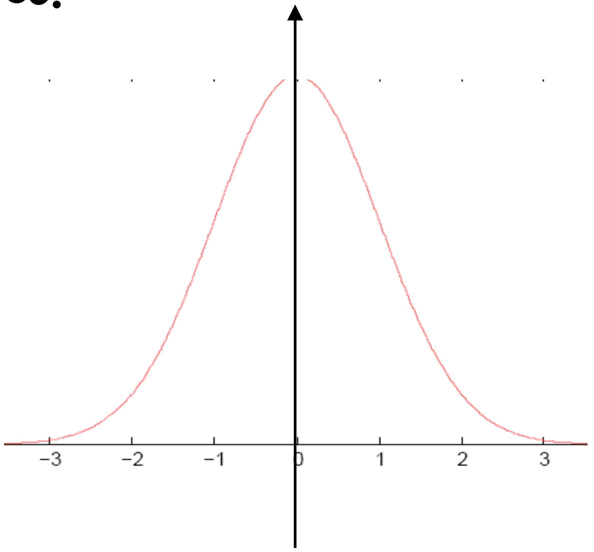
$$=x(-e^{-\lambda x})\Big|_0^{+\infty}+\int_0^{+\infty} e^{-\lambda x}dx$$

$$\int u\,dv = uv - \int v\,du$$

$$=-\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{+\infty}=\frac{1}{\lambda}.$$

3. 标准正态分布： X ~ N(0,1)

密度函数为：$f(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},-\infty<x<+\infty.$



注意到:$\int_0^{+\infty}x\,e^{-\frac{x^2}{2}}dx=-e^{-\frac{x^2}{2}}\Big|_0^{+\infty}=1,$

所以$\int_{-\infty}^{+\infty}|x|\,f(x)$收敛，从而$X$的期望存在。

$\Rightarrow\int_{-\infty}^{+\infty}xf(x)dx$亦收敛，

又因其柯西主值为0（积分$\int_{-A}^{A}xf(x)dx$的被积函数为奇函数）

$\Rightarrow E(X)=0.$

4. Cauchy分布

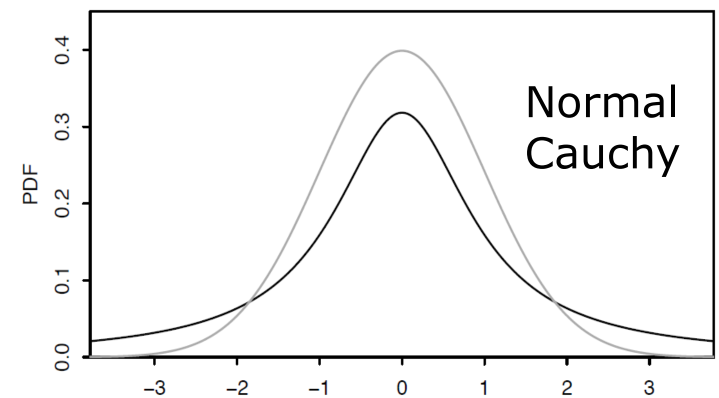
密度函数为：$f(x)=\dfrac{1}{\pi}\dfrac{1}{1+x^2},-\infty<x<+\infty.$

首先，$\int_{-\infty}^{+\infty}f(x)dx=\dfrac{1}{\pi}\arctan x\,|_{-\infty}^{+\infty}=1\Rightarrow f(x)$是密度函数。

由于$\int_{-\infty}^{+\infty}|x|\,f(x)dx=+\infty$

(事实上，$\int_{0}^{+\infty}xf(x)dx=\int_{0}^{+\infty}x\dfrac{1}{\pi}\dfrac{1}{1+x^2}dx$

$$=\frac{1}{2\pi}\int_{0}^{+\infty}\frac{d(1+x^2)}{1+x^2}dx=\frac{1}{2\pi}\ln(1+x^2)\,|_{0}^{+\infty}=+\infty)$$



所以Cauchy分布的数学期望不存在。

问：积分$\int_{-\infty}^{+\infty}xf(x)dx$收敛吗？ **No!** 虽然其柯西主值为0.

五、RV函数的数学期望

问题： 设X为一RV，Y = g(x) 为另一RV，如何求Y的期望？

例： 设**X**的分布列为：

X	**-2**	**-1**	**0**	**1**	**2**
P	**0.2**	**0.1**	**0.1**	**0.3**	**0.3**

求 **Y = X^2** 的期望。

解1：先求$Y = X^2$的分布列：

Y	$(-2)^2$	$(-1)^2$	0	1	2^2
P	0.2	0.1	0.1	0.3	0.3

合并$\Rightarrow$

Y	0	1	4
P	0.1	0.4	0.5

$$\Rightarrow E(Y) = 0\times 0.1 + 1\times 0.4 + 4\times 0.5 = 2.4.$$

解2：在合并前的表中直接算：

$$E(Y) = (-2)^2\times 0.2 + (-1)^2\times 0.1 + 0\times 0.1 + 1\times 0.3 + 2^2\times 0.3 = 2.4.$$

定理：Law Of Unconscious Statistician (LOTUS)

设$RV\ X$的分布列为$P(X=x_k)=p(x_k)$，$k=1,2,\ldots$(离散型)，

或具有概率密度函数$f(x)$(连续型)，

则X的某一函数$Y=g(X)$的数学期望为：

$$E(Y)=E[g(X)]=\begin{cases}\sum g(x_k)p(x_k), & X\text{为离散型；}\\ \int g(x)f(x)dx, & X\text{为连续型。}\end{cases}$$

对比E(X)的定义

just to change x to g(x) in the definition of E(x)

(设所涉及的数学期望均存在)

证：设$y_j,\ j=1,2,\ldots$表示$g(x_k)$，$k=1,2,\ldots$的不同取值

$$\text{则}\sum_k g(x_k)p(x_k)=\sum_j\ \sum_{i:g(x_i)=y_j} g(x_i)\ p(x_i)$$

$$=\sum_j y_j\ P(g(X)=y_j)=\sum_j y_j\ P(Y=y_j)=E(Y).$$

该定理的重要性是不需要知道 Y=g(X) 的分布，只用到X的分布即可。

Remark

- How do we correctly calculate E(g(X))? LOTUS provided an answer.

- It is possible to find E(g(X)) directly using the distribution of X, without first having to find the distribution of g(X).

- Note that in general

  E(g(X)) does not equal g(E(X))

  if g is not linear.

例:

某原材料的市场需求 X~U(300, 500)（吨）。每出售一吨，

公司获利1.5万元；积压一吨，损失0.5万元。问公司应组织多少货源，可使平均收益最大?

解：设应组织a吨，则$300 < a < 500$.

收益Y为市场需求X的函数，

$$Y = g(X) = \begin{cases} 1.5a, & X \ge a; \\ 1.5X - 0.5(a - X), & X < a. \end{cases}$$

不需要知道 Y 的分布

$$E(Y) = \int_{-\infty}^{+\infty} g(x) p_X(x) dx = \int_{300}^{500} g(x) \cdot \frac{1}{200} dx$$

$$= \int_{300}^{a} g(x) \cdot \frac{1}{200} dx + \int_{a}^{500} g(x) \cdot \frac{1}{200} dx = \frac{1}{200}(-a^2 + 900a - 300^2).$$

关于a求导，令之为0，得$a = 450$.

关于a求二阶导，小于0，所以为最大值。

五、数学期望的性质

(1) 若 C 为常数，则 $E(C) = C$.

(2) 对任意常数 a，有 $E(aX) = aE(X)$.

(3) 对任意函数 $g_1(X)$和 $g_2(X)$，有

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)].$$

特别，$E(aX+b)= aE(X)+ b$.

$$E[g(X)] = \sum_k g(k)P(X = k)$$

证:(2)在**LOTUS**定理中取$g(X)=aX$;

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

(3)在**LOTUS**定理中取$g(X)=g_1(X)+g_2(X)$;

再注意到“求和”和“积分”的**线性性**即可。

例：一般正态分布的数学期望

设$X \sim N(\mu,\sigma^2)$,求$E(X)$.

解：令$U := \dfrac{X-\mu}{\sigma}$,则$U \sim N(0,1)$.

由$U = \dfrac{X-\mu}{\sigma}$得 $X = \sigma U + \mu$,

由数学期望的线性性：

$$E(X) = E(\sigma U + \mu) = E(\sigma U) + E(\mu)$$
$$= \sigma E(U) + \mu = \mu.$$

正态分布的第一个参数为其数学期望。

六、数学期望的统计意义

设$m = E(X)$.

问：如何选取一个常数，去“最优”近似$RV\ X$?

或：$E(X-a)^2$在a为何值时最小？

$$E(X-a)^2 = E(X-m+m-a)^2$$

$$= E[(X-m)^2 + 2(X-m)(m-a) + (m-a)^2]$$

$$= E[(X-m)^2] + 2(m-a)\underline{E[X-m]} + E[(m-a)^2]$$

$$= E[(X-m)^2] + [(m-a)^2] \geq E[(X-m)^2].$$

当$a = m = E(x)$时，$E(X-a)^2$达最小，

最小值为$E[(X-m)^2]$.

第五节：方差与标准差

例：设有两股票

问：买哪种股票更理想？

比较A，B在未来的“平均”价格：

A: 110*0.7+90*0.3 = 77+27=104;

B: 120*0.6+80*0.4 = 72+32=104;

完全一样！

例：设有两股票

如何刻画“风险”？

——“波动”的大小！即偏离“均值”的程度！

A: 0.7*(110-104)^2 + 0.3*(90-104)^2 = 84;

B: 0.6*(120-104)^2 + 0.4*(80-104)^2 = 384.

所以，A的波动（风险）更小，应优先考虑。

一、方差的定义

定义：设 X 为一RV,若 $E(X^2)$ 存在，则称

$$\textbf{Var}(X) := D(X) := E(X-EX)^2$$

为 X 的方差。称$\sigma(X) := \sqrt{D(X)}$ 为 X 的标准差。

注记：

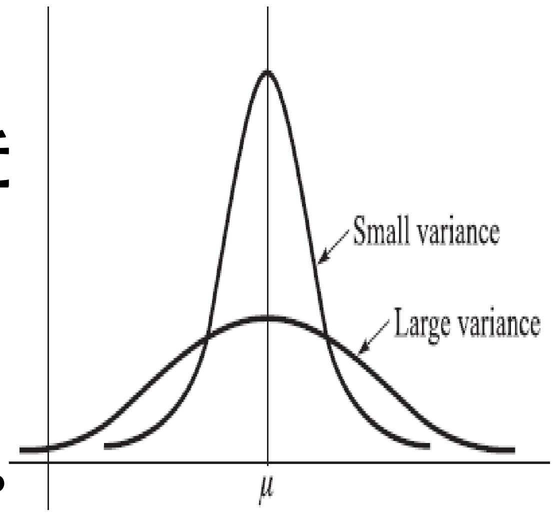
(1) 方差反映了RV相对其均值的偏离程度。

方差(标准差)小，表明RV集中在E(X)附近

方差(标准差)大，表明RV分散；

(2) 方差与标准差量纲不同。

标准差具有与RV本身及E(x)相同的量纲。



注记：

(3) $\mathbf{Var}(X) := D(X) := E(X-EX)^2$

由LOTUS公式

$$
= \begin{cases} \sum (x_k - EX)^2 p_k, & \text{离散型} \\ \int (x-EX)^2 f(x)dx, & \text{连续型} \end{cases}
$$

Variance of *X* is the expected value of a specific function of *X*, g(x) = (x-EX)^2.

(4) 若$E(X^2)$存在，则$E(X)$存在（因$|x| \le x^2+1$);

且$E(X-EX)^2$也存在

（因$E(X-EX)^2 = E(X^2) - [E(X)]^2$).

二、方差的性质

(1) $D(X)=E(X^2)-[E(X)]^2$.

证：$D(X)=E[X-E(X)]^2=E\{X^2-2XE(X)+[E(X)]^2\}$

$=E(X^2)-2E(X)\cdot E(X)+[E(X)]^2=E(X^2)-[E(X)]^2$.

(2) 常数的方差为0，即$D(C)=0$

(反之，如果$D(X)=0$，则对X有何判断?)

(3) 若a,b为常数，则$D(aX+b)=a^2D(X)$.

$D(aX+b)=E[(aX+b)-E(aX+b)]^2$

$=E[a\ (X-EX)]^2=a^2D(X)$.

二、方差的性质

随机变量的标准化

设$D(X)>0$，　令$Y=\dfrac{X-E(X)}{\sqrt{D(X)}}$,

则$E(Y)=0, D(Y)=1.$

称Y为X的标准化。

三、常见RV的方差

1. 两点分布

X	0	1
P	1-p	p

E(x)=p, E(X^2)=p, D(X) = p (1-p)

2. 二项分布

设$X \sim B(n,p)$,即$P(X=k)=C_n^k p^k (1-p)^{n-k}, k=0,1,\ldots,n.$

则$E(X)=np,$

$$E(X^2)=\sum_{k=0}^{n} k^2 C_n^k p^k (1-p)^{n-k} = \ldots = n(n-1)p^2+np,$$

$$D(X)=np(1-p).$$

3. 几何分布

设$X \sim Ge(p)$,即$P(X=k)=p(1-p)^{k-1}, k=1,2,\ldots$

则$E(X)=\sum_{k=1}^{\infty} kP(X=k)=\dfrac{1}{p}$.

$$E(X^2)=\sum_{k=1}^{\infty} k^2 P(X=k)=\ldots=\frac{2(1-p)}{p^2}+\frac{1}{p},$$

$$D(X)=\frac{(1-p)}{p^2}.$$

4. Poisson分布

设$X \sim P(\lambda)$,即$P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\ldots$

则$E(X)=\sum\limits_{k=0}^{\infty} k\, P(X=k)=\lambda,$

$$E(X^2)=\sum_{k=0}^{\infty} k^2\, P(X=k)=\lambda^2+\lambda.$$

$\Rightarrow D(X)=E(X^2)-[E(X)]^2=\lambda.$ 参数

(Poisson分布的均值、方差均为λ)

5. 均匀分布

设$X \sim U(a,b)$，即概率密度函数为

$$f(x)=\frac{1}{b-a}I_{\{a<x<b\}}.$$

则 $$E(X)=\frac{a+b}{2},$$

$$E(X^2)=\int_a^b \frac{x^2}{b-a}dx=\frac{1}{3}(a^2+ab+b^2)$$

$$\Rightarrow D(X)=\frac{1}{12}(b-a)^2.$$

6. 指数分布

设$X\sim \mathbf{Exp}(\lambda)$，即$X$的概率密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0,\\ 0, & \text{其它},\end{cases}$$

则 $E(X)=1/\lambda,$

$$E(X^2)=\int_0^{+\infty} x^2\lambda e^{-\lambda x}dx=\int_0^{+\infty} x^2 d(-e^{-\lambda x})$$

$$=x^2(-e^{-\lambda x})\big|_0^{+\infty}+\int_0^{+\infty} 2xe^{-\lambda x}dx$$

Integration by part

$$\int u\,dv=uv-\int v\,du$$

$$=2\frac{1}{\lambda}\underline{\int_0^{+\infty} x\ \lambda e^{-\lambda x}dx}=\frac{2}{\lambda^2}.$$

$$\Rightarrow D(X)=E(X^2)-[E(X)]^2=1/\lambda^2.$$

7. 正态分布

标准正态密度函数为：

$$f(u)=\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}},-\infty<u<+\infty.$$

已知 $E(U)=0,$

$$E(U^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}u^2\ e^{-\frac{u^2}{2}}du=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}u\ d(-e^{-\frac{u^2}{2}})$$

$$=\frac{1}{\sqrt{2\pi}}\{-ue^{-\frac{u^2}{2}}\big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}e^{-\frac{u^2}{2}}du\}$$

Integration by part

$$\int u\,dv=uv-\int v\,du$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{u^2}{2}}du=1\ .$$

$$\Rightarrow\quad D(U)=1.$$

7. 正态分布

对一般正态分布$X \sim N(\mu, \sigma^2)$,

令$U := (X-\mu)/\sigma$,则$U \sim N(0,1)$

注意到 $X = \sigma U + \mu$

$\Rightarrow D(X) = \sigma^2 D(U) = \sigma^2.$

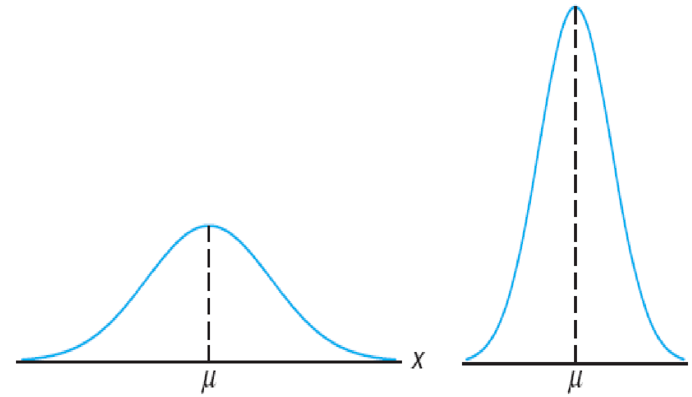
正态分布的第二个参数σ^2为方差。

四、切比雪夫不等式

问题：

$E(X): RV\ X$ 集中的位置;

$D(X): RV\ X$ 偏离“中心”的程度；

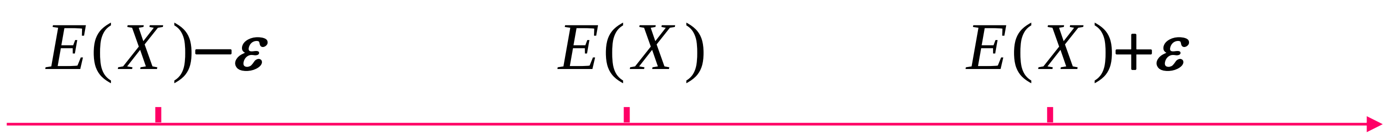


(a) (b)

问：事件$\{|X-E(X)<\varepsilon|\}$,

即 X 落在$(E(X)-\varepsilon, E(X)+\varepsilon)$ 的概率有多大？

或问：$\{|X-E(X)\geq\varepsilon|\}$ (较大偏差)的概率有多大？



切比雪夫不等式在“大数定律”中有重要应用。

定理： 设RV X有数学期望$E(X)=\mu$,有方差$D(X)=\sigma^2$.

则$\forall\,\varepsilon>0$,有$P(|X-E(X)|\geq\varepsilon)\leq\dfrac{\sigma^2}{\varepsilon^2}$,（方差小，有大偏差的概率小）

等价地，有$P(|X-E(X)|<\varepsilon)\geq 1-\dfrac{\sigma^2}{\varepsilon^2}$.

证：(只对连续型情形证明)

$$P(|X-\mu|\geq\varepsilon)=\int_{|x-\mu|\geq\varepsilon} f(x)dx$$

$$\leq\int_{|x-\mu|\geq\varepsilon}\frac{(x-\mu)^2}{\varepsilon^2}f(x)dx \quad \text{(放大被积函数)}$$

$$\leq\int_{-\infty}^{+\infty}\frac{(x-\mu)^2}{\varepsilon^2}f(x)dx \quad \text{(放大积分区域)}$$

$$=\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}(x-\mu)^2 f(x)dx=\frac{\sigma^2}{\varepsilon^2}.$$

Remarks

- Chebyshev's inequality enable us to derive bounds on probability when only mean and variance are known.

- Chebyshev's inequality holds for any distribution and thus the results are usually weak.

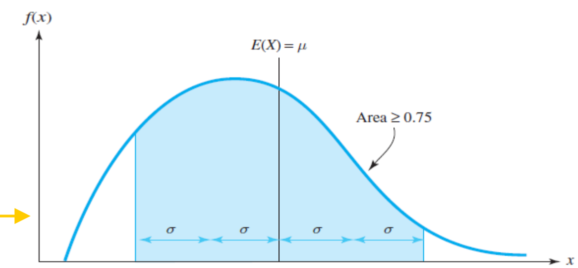
- The use of Chebyshev's inequality is relegated to situations where the distribution is unknown.

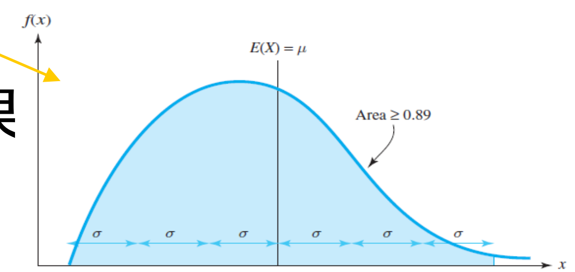
注记：

切比雪夫不等式给出了在 RV 分布未知，而只知道 E(X) 和 D(x) 的情况下的概率估计。这种估计一般比较**粗糙**。

如果设$E(X)=\mu, D(X)=\sigma^2$,则由切比雪夫不等式，有(对任何分布成立)

$$P(|X-\mu|\le k\sigma)\ge 1-\frac{1}{k^2}=\begin{cases}0, & k=1;\\ 3/4, & k=2;\\ 8/9, & k=3.\end{cases}$$





对比：如果X的分布已知,则有精确结果

如设$X\sim N(\mu,\sigma^2)$,

$$\text{则 } P(|X-\mu|\le k\sigma)=2\Phi(k)-1=\begin{cases}0.6826, & \text{当 } k=1;\\ 0.9545, & \text{当 } k=2;\\ 0.9973, & \text{当 } k=3.\end{cases}$$

定理： 设**RV** X的方差存在，则$D(X)=0 \Leftrightarrow P(X=a)=1$

"X几乎处处为某常数".

证：充分性显然，下证必要性。

设$D(X)=0$，则$E(X)$存在。

$$\{|X-EX|>0\}=\bigcup_{n=1}^{+\infty}\{|X-EX|>1/n\}$$

单调增事件序列的极限事件

$$0\le P(\{|X-EX|>0\})=P(\bigcup_{n=1}^{+\infty}\{|X-EX|>1/n\})$$

次可加性

$$\le \sum_{n=1}^{+\infty} P(\{|X-EX|>1/n\}) \le \sum_{n=1}^{+\infty} \frac{D(X)}{1/n^2}=0,$$

切比雪夫不等式

$$\Rightarrow P(\{|X-EX|>0\})=0,$$

则对立事件$P(\{|X-EX|=0\})=1 \Rightarrow P(X=EX)=1.$

第六节：分布的其它特征数

一、k阶矩

k 阶原点矩：$\mu_k := E(X^k)$;

k 阶中心矩：$\nu_k := E[(X-EX)^k]$;

特例：一原点矩：$\mu_1 := E(X)$ ———数学期望；

二阶中心矩：$\nu_2 := E[(X-EX)^2]$———方差。

注：若k阶矩存在，则更低阶的矩都存在

(因$|x|^{k-1} \le |x|^k + 1, \forall x$)

中心矩与原点矩的关系：k阶中心矩

$$\nu_k = E[(X-\mu_1)^k] = \sum_{i=0}^{k} C_k^i EX^i(-\mu_1)^{k-i} = \sum_{i=0}^{k} C_k^i \mu_i (-\mu_1)^{k-i}.$$

例： 设$X \sim N(0,\sigma^2)$. 求$E(X^k)$.

解：$|x|^k e^{-\frac{x^2}{2\sigma^2}} = |x|^k e^{-\frac{x^2}{4\sigma^2}} e^{-\frac{x^2}{4\sigma^2}} \le \mathbf{C}\, e^{-\frac{x^2}{4\sigma^2}}$ （存在这样的常数C）

$$\int_{-\infty}^{+\infty} |x|^k e^{-\frac{x^2}{2\sigma^2}} dx \le \mathbf{C} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{4\sigma^2}} dx < \infty.$$

所以，$E(X^k) = \dfrac{1}{\sqrt{2\pi}\,\sigma} \displaystyle\int_{-\infty}^{+\infty} x^k e^{-\frac{x^2}{2\sigma^2}} dx$存在。

当k为奇数时，$E(X^k) = 0$;

当k为偶数时，$E(X^k) = \dfrac{\sigma^2}{\sqrt{2\pi}\,\sigma} \displaystyle\int_{-\infty}^{+\infty} x^{k-1} d\left(-e^{-\frac{x^2}{2\sigma^2}}\right)$

Integration by parts

$$\int u\,dv = uv - \int v\,du$$

$$= (k-1)\sigma^2 \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} x^{k-2} e^{-\frac{x^2}{2\sigma^2}} dx = \sigma^2 (k-1) E(X^{k-2})$$

$$= \sigma^4 (k-1)(k-3) E(X^{k-4}) = \sigma^{k-2}(k-1)\cdots 3\, E(X^2) = \sigma^k (k-1)!!$$

二、变异系数

问题：如何比较量纲不同的两个随机变量的波动大小？

定义：设随机变量的方差存在，则称比值

$$C_v(X)=\frac{\sqrt{D(X)}}{E(X)}=\frac{\sigma(X)}{E(X)}$$

为随机变量的变异系数。(注：$C_v(X)$是无量纲的量，消除量纲的影响)

例： 同龄树高度 X（米）：E(X)=10，D(X)=1；

某年龄段儿童身高 Y（米）：E(Y)=1，D(Y)=0.04；

谁的波动大？

直接比较方差？ NO

比较变异系数： $C_v(X)=1/10=0.1$

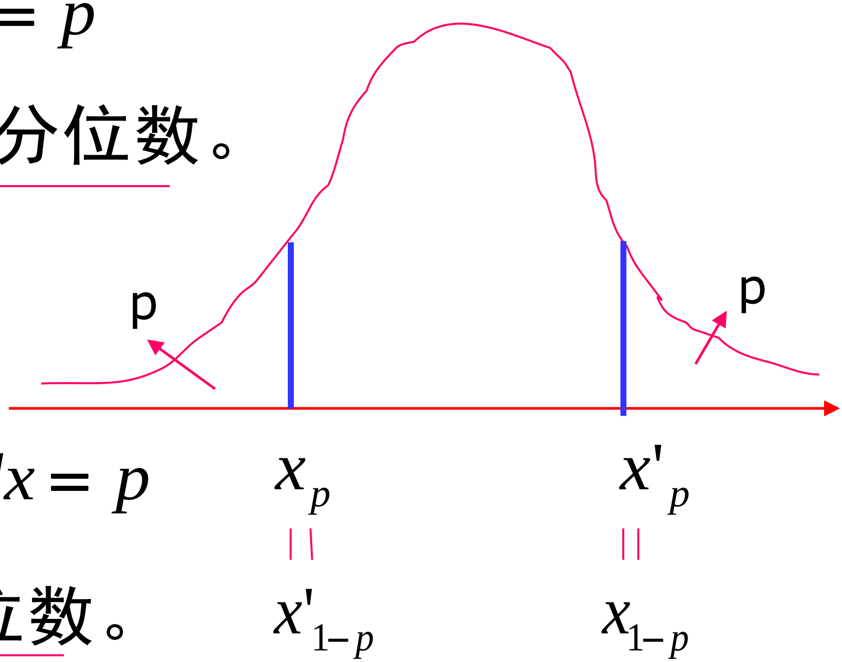
$C_v(Y)=0.2/1=0.2$，Y的波动更大。

三、分位数（quantile）

定义: 设连续型RV X的密度函数为$f(x)$,分布函数为$F(x)$,对任意$p \in (0,1)$,称满足条件

$$F(x_p) = \int_{-\infty}^{x_p} f(x)dx = p$$

的数 x_p 为此分布的(下侧) p分位数。



同理，称满足条件

$$1 - F(x'_p) = \int_{x'_p}^{+\infty} f(x)dx = p$$

数 x'_p 为此分布的上侧p分位数。

关系：$x'_{1-p} = x_p$， $x_{1-p} = x'_p$.

三、分位数（quantile）

对一般分布：

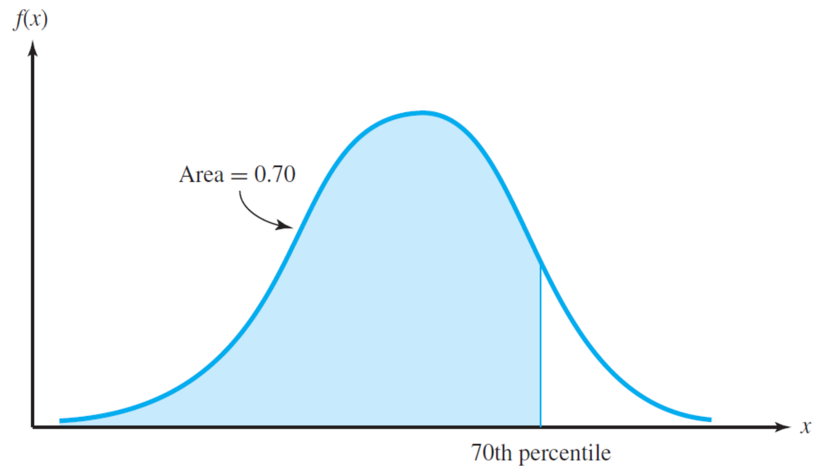
> **Definition** . For $0 < p < 1$, the pth quantile of a random variable X is any real value x satisfying
>
> $$P(X \geq x) \geq 1 - p \quad \text{and} \quad P(X \leq x) \geq p.$$

X左边（含x），概率不小于p
X右边（含x），概率不小于（1-p）

三、分位数（百分位数 Percentile)

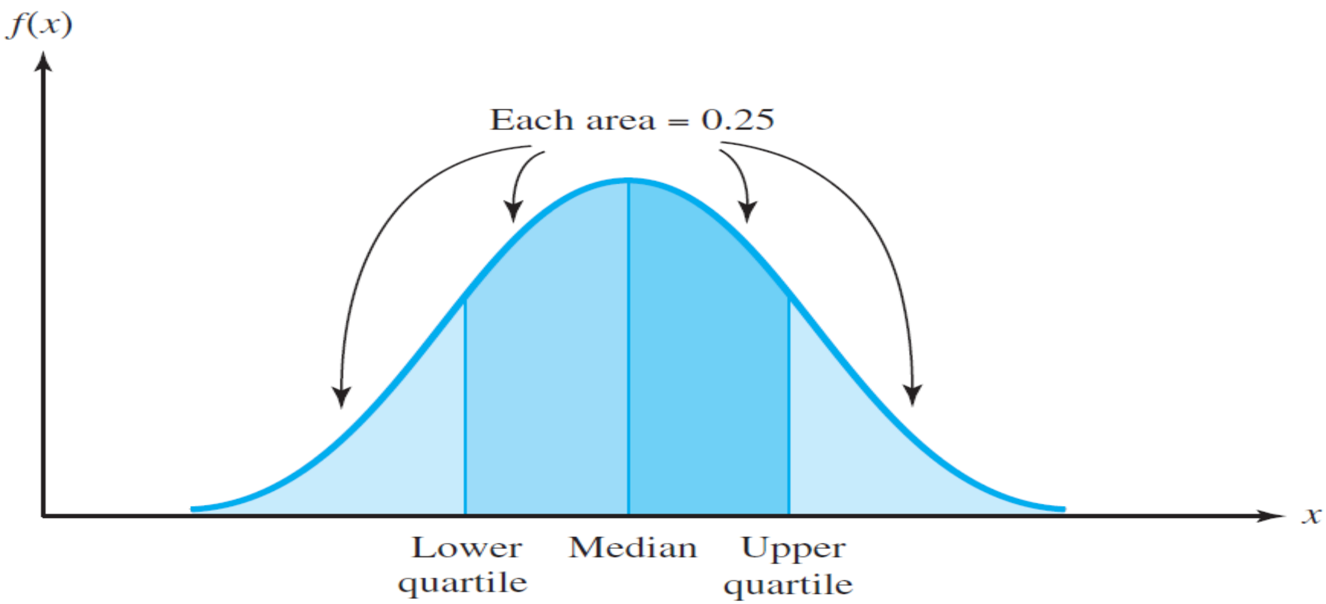
- The prob *p* is often written as a percentage, and resulting quantiles are called percentiles.
- For instance, the 70th percentile of distribution is the value *x* for which

$$F(x) = 0.70.$$



- The quantiles corresponding to p = 0.25 and p = 0.75 are called the first and third quartiles.

Remark：Quartile（四分位数）



下四分位数
25^{th} percentile.

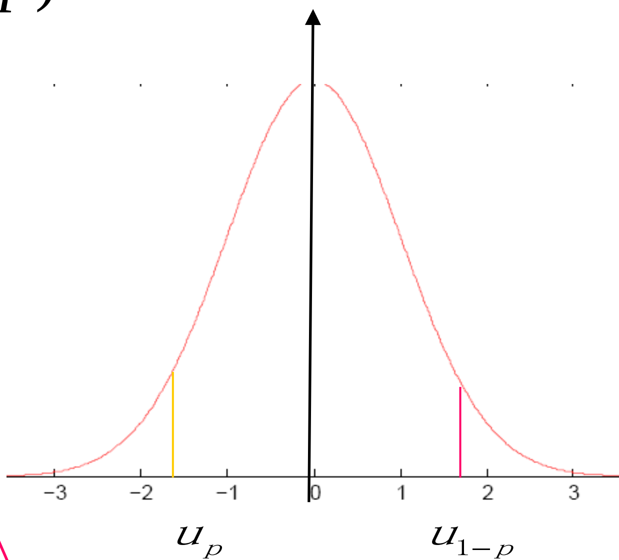
上四分位数
75^{th} percentile.

Interquartile range = upper quartiles – lower quartile.
It provides an indication of the spread of the distribution.

例：正态分布的分位数

给定$0 < p < 1$,标准正态分布的下侧分位数u_p 满足

$$\Phi(u_p) = \int_{-\infty}^{u_p} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = p \Rightarrow u_p = \Phi^{-1}(p).$$



比如：查表$u_{0.95} = 1.645$.

性质：当$p < 0.5$时，$u_p < 0$;

当$p > 0.5$时，$u_p > 0$;

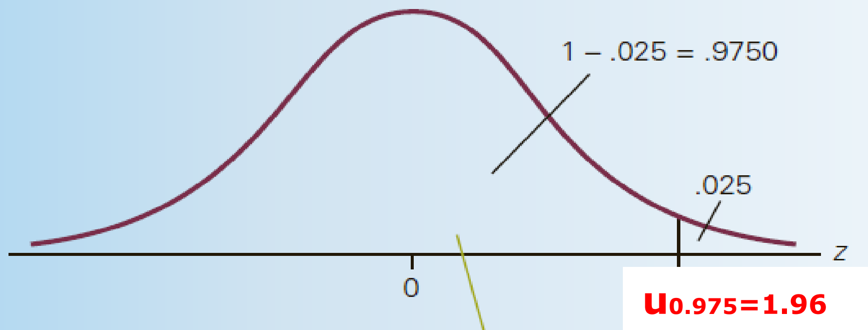
当$p = 0.5$时，$u_p = 0$;

对任意p, $u_p = -u_{1-p}$.

一般正态分布$N(\mu, \sigma^2)$的下侧分位数$x_p = \sigma u_p + \mu$,因为

$$p = \int_{-\infty}^{x_p} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\frac{x_p-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \Rightarrow \frac{x_p - \mu}{\sigma} = u_p.$$

查表：下侧0.975（上侧0.025）分位点



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916

R commands

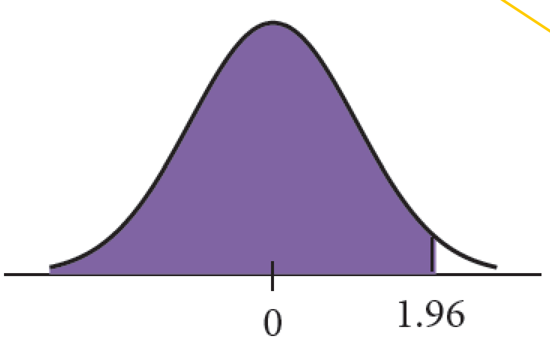
R Commands for the Normal(μ, σ^2) Distribution

```
dnorm(x,μ,σ) # gives the PDF for x in (−∞,∞)

pnorm(x,μ,σ) # gives the CDF for x in (−∞,∞)

qnorm(s,μ,σ) # gives the s100th percentile for s in (0,1)

rnorm(n,μ,σ) # gives a sample of n normal(μ,σ²) random
  variables
```



qnorm(0.95, 0, 1)=1.96

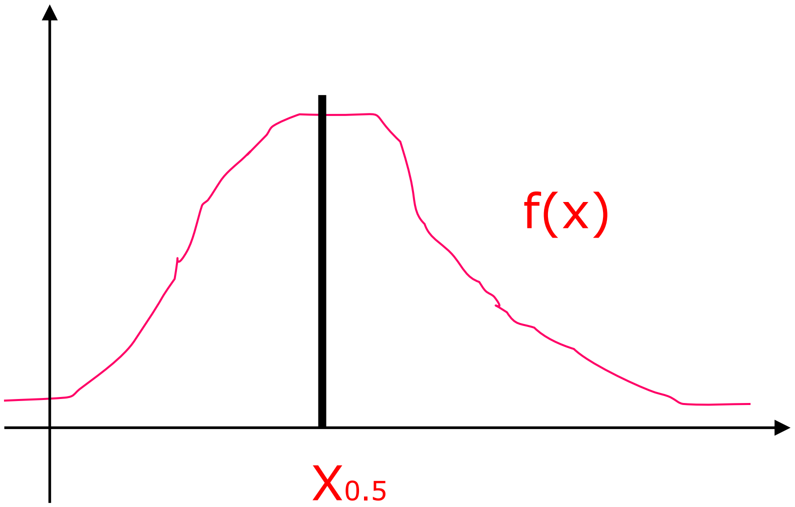
四、中位数 (Median)

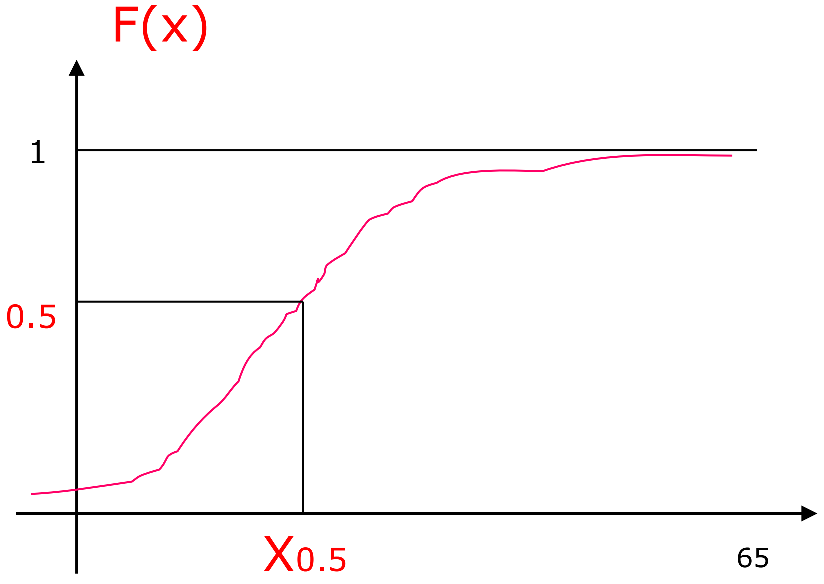
对应于 p=0.5 时的分位数是中位数。

设连续型$RV\ X$的密度函数为$f(x)$,分布函数为$F(x)$,称满足条件

$$F(x_{0.5})=\int_{-\infty}^{x_{0.5}} f(x)dx=0.5$$

的数 $x_{0.5}$ 为此分布的中位数。





注记：

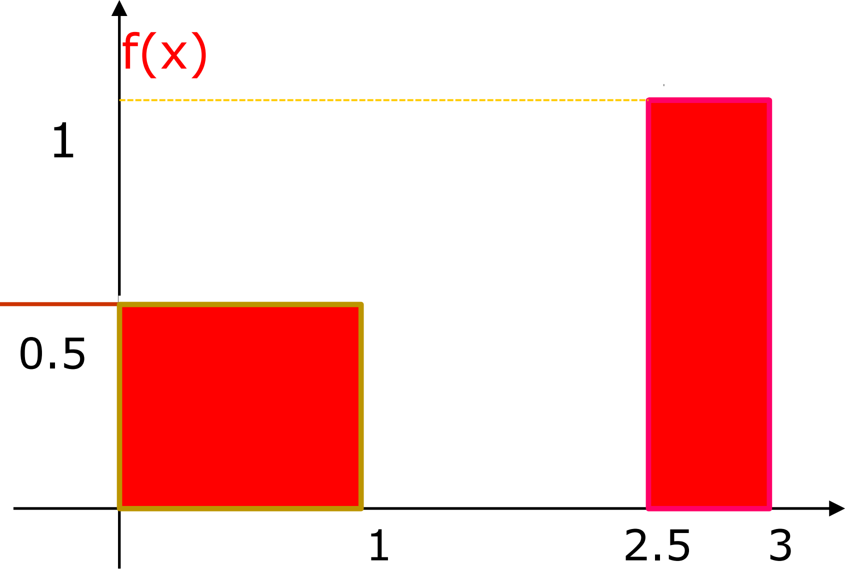
- **中位数可能不唯一。**

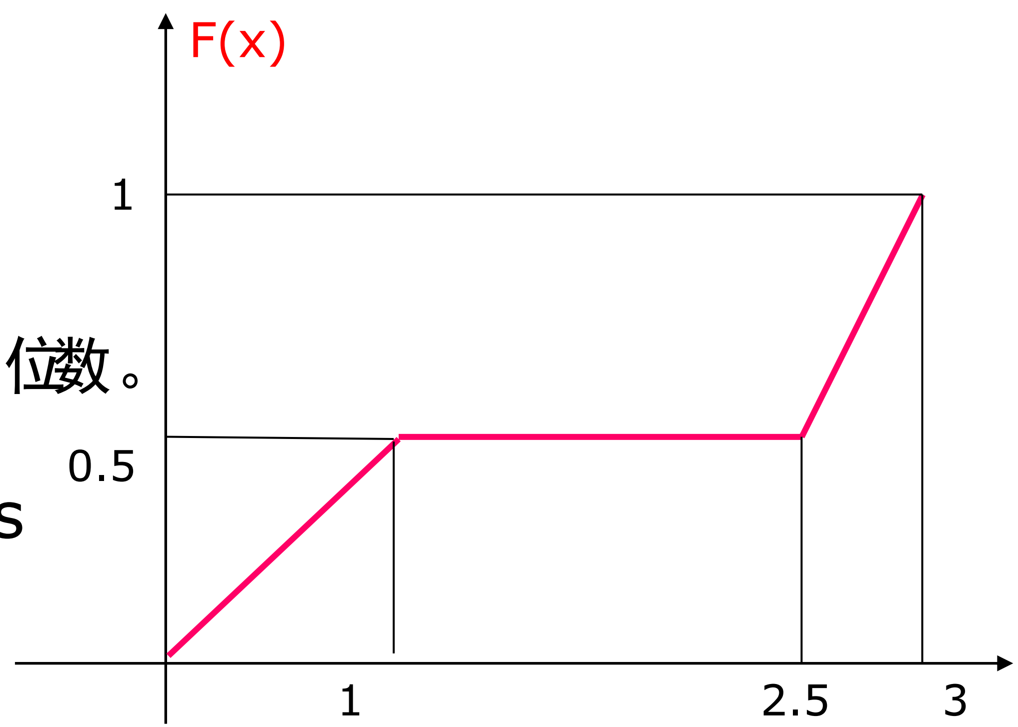
考虑连续型$RV\ X$,其密度函数为

$$f(x)=\begin{cases}1/2, & x\in[0,1];\\ 1, & x\in[2.5,3];\\ 0, & 其它.\end{cases}$$

对该分布，

[1, 2.5]之间的任何数都是中位数。





When CDF is continuous and strictly increasing, the median is unique.

一般分布的中位数

数m称为中位数，如果 $P(X \overset{\text{含m}}{\le} m) \ge \frac{1}{2}$, $P(X \overset{\text{含m}}{\ge} m) \ge \frac{1}{2}$.

m左、右边(含)的概率均至少为0.5.

例：

X	1	2	3	4
P	0.1	0.2	0.3	0.4

$P(X \le 3) = 0.6, P(X \ge 3) = 0.7 \Rightarrow$ 中位数$m = 3$(唯一)。

例：

X	1	2	3	4
P	0.1	0.4	0.3	0.2

$P(X \le 2) = 0.5, P(X \ge 3) = 0.5$

$\Rightarrow [2,3]$中的任何数均为中位数(不唯一)。

中位数与均值：谁是更好的“平均”的度量？

1） 都是反映随机变量的位置特征。
2） 任何分布的中位数都存在，但均值(期望)不一定；
3） 均值(期望)对“大的取值”上的概率变化很敏感，
但中位数几乎不受影响（考虑某富豪搬入某村对均值和中位数的影响）

比较：

- 全村100个家庭“平均”年收入 30 万元
（可能是99个家庭年收入 1万，另一家庭2901万）.
- 全村100个家庭年收入“中位数”为 30 万元：
至少有一半家庭的年收入达到或超过 30 万元。

在许多场合，
“中位数”是比“均值”(期望)更有用的“平均”的度量。

中位数与均值的比较

- 均值 $E(X)$

$$E(X - EX)^2 = \min_d E(X - d)^2;$$

--- minimizing the **mean squared error**;

- 中位数 m

$$E|X - m| = \min_d E|X - d|;$$

--- minimizing the **mean absolute error**.

（Please try to prove this）

五 偏度系数和峰度系数(Skewness and Kurtosis)

- 如何刻画分布偏离对称性的程度？

定义：设随机变量 X 的前三阶矩存在，则称

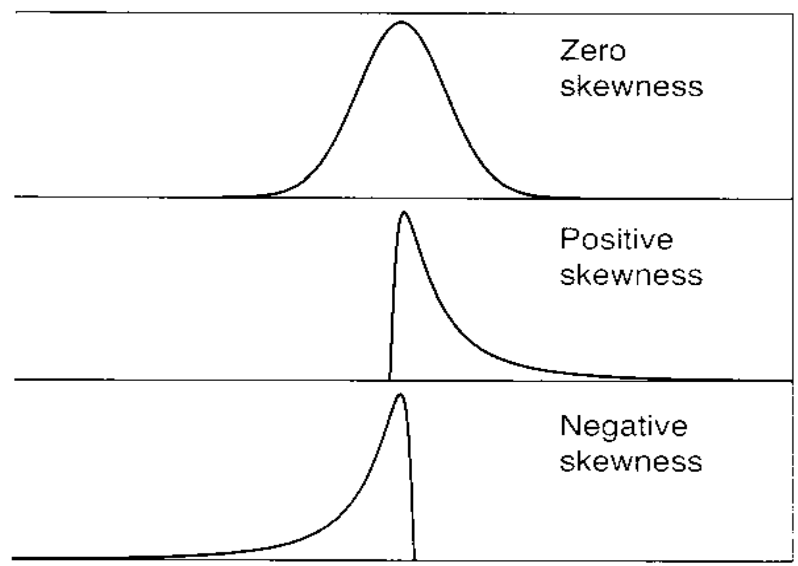
$$\beta_s(X) = E\left(\frac{X-\mu}{\sigma}\right)^3$$

以正态分布为基准

为随机变量的偏度系数(注：是无量纲的量)

$\beta_s(X) > 0$：正偏（右偏）

$\beta_s(X) < 0$：负偏（左偏）



五 偏度系数和峰度系数(Skewness and Kurtosis)

- 如何刻画分布的尖峭程度和尾部粗细程度？

定义：设随机变量 X 的前四阶矩存在，则称

$$\beta_k(X)=E\left(\frac{X-\mu}{\sigma}\right)^4-3$$

以正态分布为基准

为随机变量的峰度系数。(注：是无量纲的量)

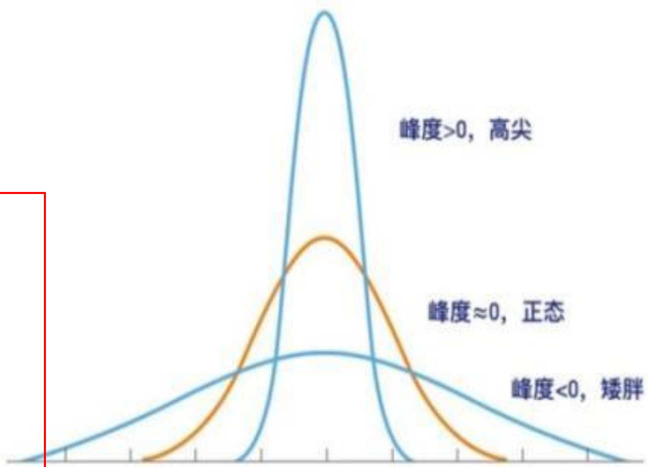
$\beta_k(X)>0$: 比 $N(0,1)$ 更尖峭, or 尾部更粗

$\beta_k(X)<0$: 比 $N(0,1)$ 更平坦, or 尾部更细

Comparison to Gaussian:

Positive: heavy-tailed (exp, gamma)

Negative: light-tailed (uniform)



第七节：随机变量的函数的分布

设 X 为一RV， $y = g(x)$ 为一函数，$Y = g(X)$ 为另一RV。

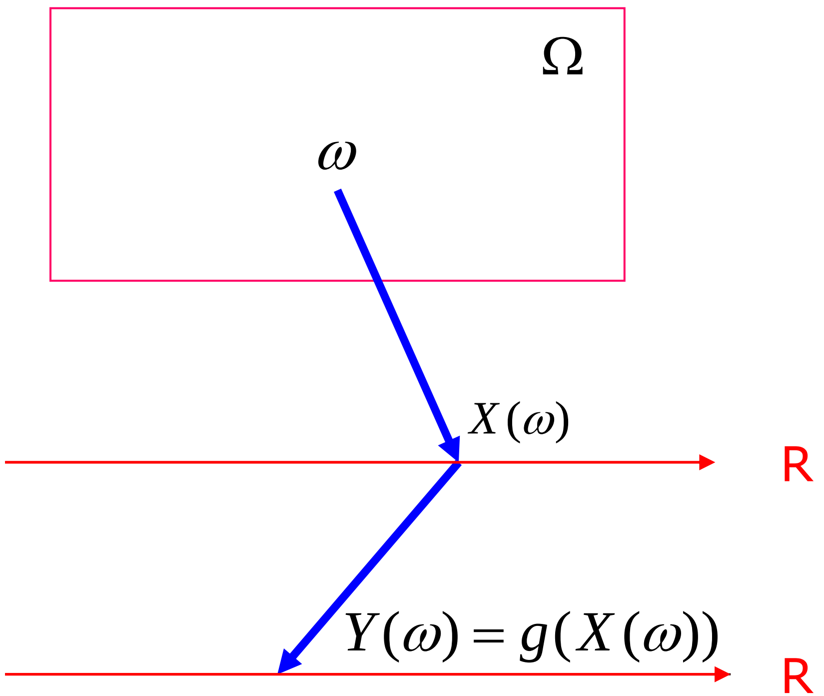
问题：求 RV Y 的分布。

如： Y = |X| ,

Y = X^2,

Y = sin(X)

…



一、离散型 RV 的函数的分布

- **设有离散型RV,**

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p	p_1	p_2	$\cdots$	p_n	$\cdots$

则$Y=g(X)$亦为离散型RV,且

$Y=g(X)$	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_n)$	$\cdots$
p	p_1	p_2	$\cdots$	p_n	$\cdots$

合并 $g(x_i)$中相等的值，并把对应概率相加

即得 Y 的分布列：

$$P(Y=y_j)=\sum_{i:g(x_i)=y_j} P(X=x_i).$$

将$g(x_i)$ 一一列出，再将相等的值合并.

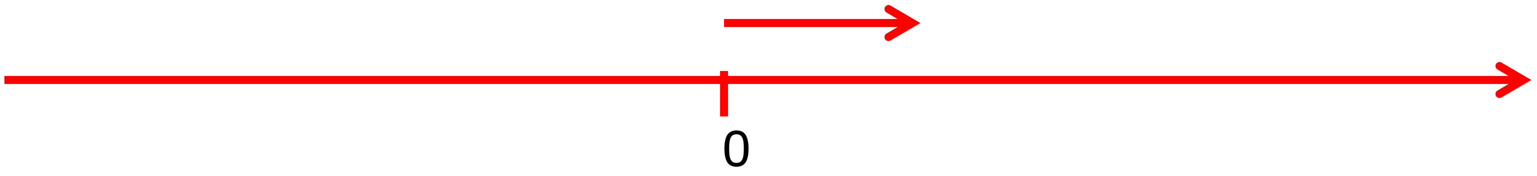
Example: Random Walk

A particle moves n steps on a number line.

It starts at 0, and at each step it moves 1 unit to the right or to the left with equal prob.

Assume all steps are independent.

Let Y be the particle's **position** after n steps. Find the distribution of Y.



Example: Random Walk

Consider each step to be a Bernoulli trial, where right is consider a success.

Let X to be the number of success, then

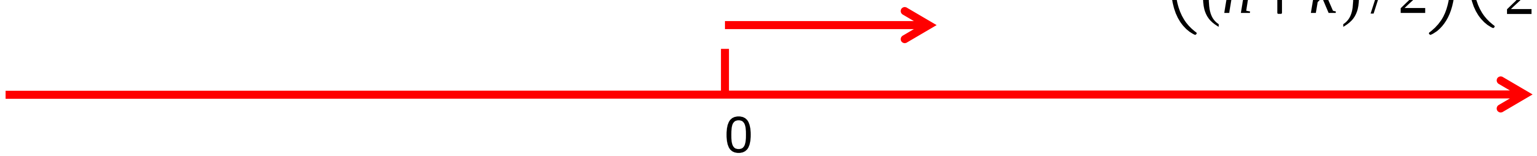
$$X \sim B(n,1/2).$$

If X=j, then the particle has taken j step to the right and (n-j) step to the left. The **position** is

$$Y=X-(n-X)=2X-n=2j-n.$$

X takes values in {0, 1, 2, …, n}, Y takes values in S:={-n, 2-n, 4-n, …, n}. For an integer k in S, then

$$P(Y=k) = P(2X-n=k)=P(X=(n+k)/2)=\binom{n}{(n+k)/2}\left(\frac{1}{2}\right)^n.$$



Example: Random Walk

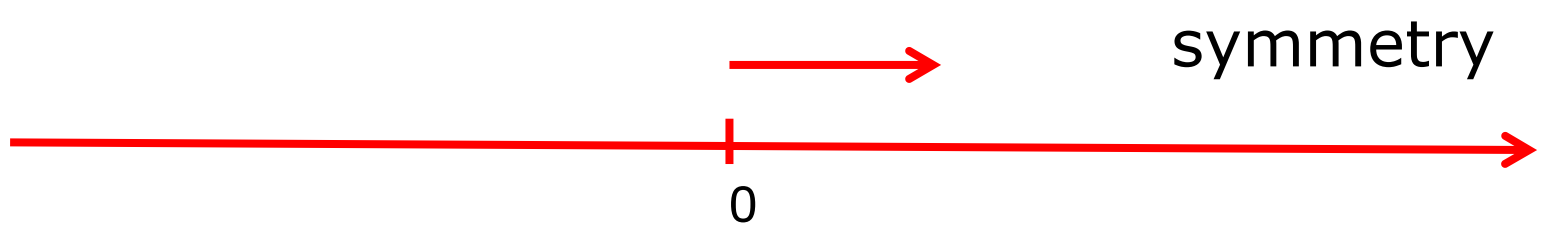
Let D be the particle's **distance** from the origin after n steps. Assume n is even. Find the distribution of D.

Note that D=|Y|. This is not one-to-one.

P(D=0) = P(Y=0) = $\binom{n}{n/2}\left(\frac{1}{2}\right)^n.$

For k=2, 4, …, n,

P(D=k) = P(Y=k) + P(Y=-k)=2 $\binom{n}{(n+k)/2}\left(\frac{1}{2}\right)^n.$



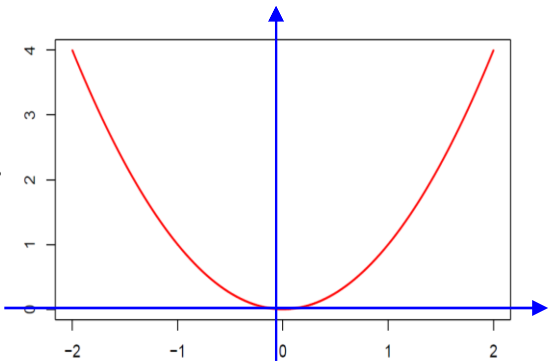
二、连续型RV的函数的分布

- Suppose that X has density function.
  Consider $Y = g(X)$ for some function g.
- What can we say about the distribution of Y ?
- The answer can be anything：
  - Y can be discrete (for an extreme case take $g(x) = 0$),
  - Y can be continuous (ex. $g(x) = x$ gives $Y = X$)
  - Y can be neither.

1. 直接法（分布函数法）

设连续型$RV\ \ X$具有概率密度函数$f_X(x)$.求$Y=X^2$的概率密度函数。

解:先求Y的分布函数。



当$y<0$时，$F_Y(y)=P(Y\le y)=P(X^2\le y)=0.$

当$y\ge 0$时，$F_Y(y)=P(Y\le y)=P(X^2\le y)$

$=P(-\sqrt{y}\le X\le\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y}).$

$$\text{对}y\text{求导}, f_Y(y)=\begin{cases}0, & y<0;\\ \dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})], & y\ge 0.\end{cases}$$

复合函数求导：
Chain rule

$$\text{特例：若}X\sim N(0,1), Y=X^2\text{，则 }f_Y(y)=\begin{cases}0, & y\le 0;\\ \dfrac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2}, & y>0.\end{cases}$$

chi-square distribution with *one degree of freedom*

例： Given an arbitrary continuous distribution, we can create a uniform distribution U(0,1).

设**RV** $X \sim F_X(x)$. 设$F_X(x)$为严格单调增的连续函数.

令$Y := F_X(X)$, 则 $Y \sim U(0,1)$.

解：下求$Y = F_X(X)$的分布函数。

不可能事件

当$y<0$时， $F_Y(y) = P(Y \le y) = P(F_X(X) \le y) = \mathbf{0};$

当$\mathbf{0} \le y < \mathbf{1}$时， $F_Y(y) = P(Y \le y) = P(F_X(X) \le y)$

$$= P(X \le F_{\mathbf{X}}^{-1}(y)) = F_X(F_{\mathbf{X}}^{-1}(y)) = y.$$

当$y \ge 1$时， $F_Y(y) = P(Y \le y) = P(F_X(X) \le y) = 1.$

必然事件

所以，**Y**的分布函数与**U(0,1)**的一致，

即$Y \sim$ **U(0,1).**

任一 RV 都可以通过其分布函数与均匀分布发生联系

例：产生任一分布随机数的方法

设$U \sim U(0,1), Y := \Phi^{-1}(U), \Phi$为 $N(0,1)$ 的分布函数。

求Y的分布函数。

解：$F_Y(y) = P(Y \le y) = P(\Phi^{-1}(U) \le y)$

$$= P(U \le \Phi(y)) = \Phi(y) \ (\text{因} U \sim U(0,1)).$$

即Y服从正态分布。

一般地，如果$U \sim U(0,1), F(x)$为**RV** X 的分布函数，

则 $F^{-1}(U) \sim F(x)$.

反函数法

Given U(0,1), we can construct an arbitrary distribution from a uniform distribution.

Theorem (Universality of Uniform)

(1) 如果$U \sim U(0,1)$, $F(x)$为某**RV** 的分布函数,

令$X := F^{-1}(U)$,则 $X \sim F(x)$.

(2) 设 $X \sim F(x)$, 则$F(X) \sim U(0,1)$.

Generate samples from distribution F(x)

- **Algorithm:**

(1) Generate $U \sim U(0,1);$

(2) Return $X = F^{-1}(U).$

(or set F(X) = U, and solve for X)

Example: (Exponential distribution)

Let X has density

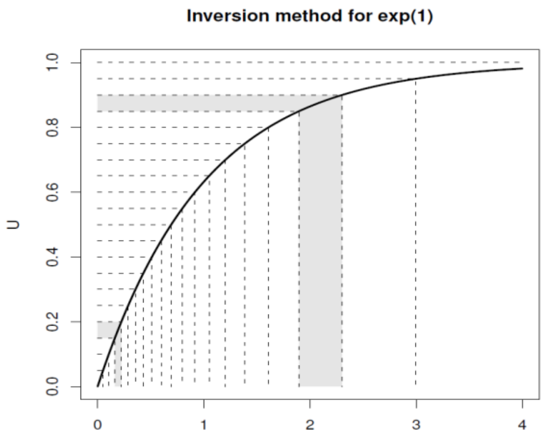
$$f(x) = \lambda e^{-\lambda x},\ x \geq 0.\ E(X) = \frac{1}{\lambda}.$$

Its distribution function is

$$F(x) = 1 - e^{-\lambda x},\ x \geq 0.$$

Then

$$X = F^{-1}(U) = -\frac{1}{\lambda}\log(1-U).$$



Note U and 1-U has the same distribution.
So we may let

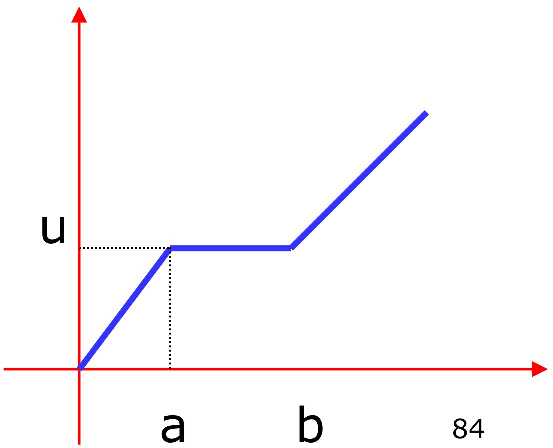
$$X = F^{-1}(U) = -\frac{1}{\lambda}\log(U).$$

Remark*

- The inverse of F is well defined when F is strictly increasing.
- Otherwise, we may need a rule to break the ties. For example, set

$$X = F^{-1}(U) := \inf\{X : F(X) \ge U\}.$$

If there are many values of x for which F(x)=u, the definition chooses the smallest.



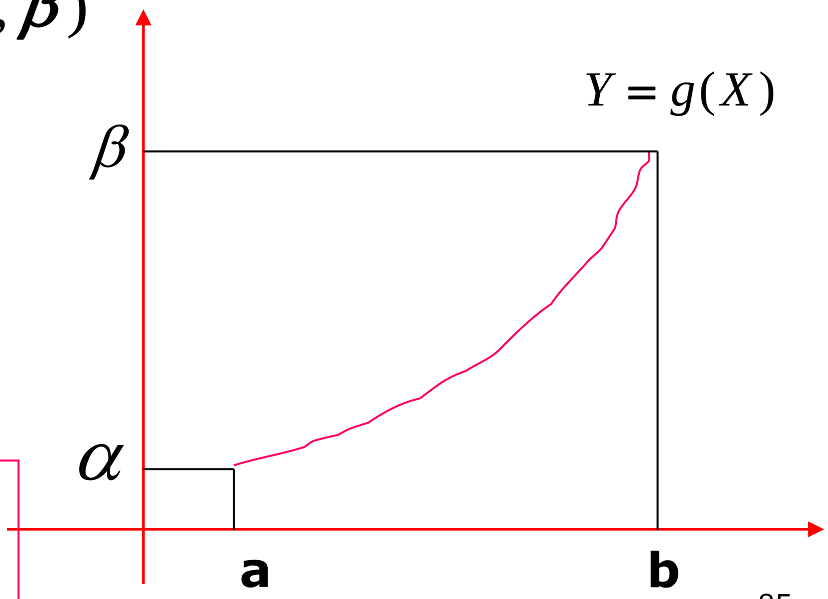
2. 变换法：定理（Jacobian Formula）

设连续型$RV\ X$具有密度 $f_X(x), x \in (a,b)$ $(a,b$可为无穷)。

设函数 $y=g(x)$处处可导且(不妨设)恒有 $g'(x)>0$.

则$Y := g(X)$为连续型RV,其概率密度函数为

$$f_Y(y)=\begin{cases} f_X(h(y))\cdot h'(y), y \in (\boldsymbol{\alpha}, \boldsymbol{\beta}) \\ 0, \qquad\qquad \text{其它.} \end{cases}$$



其中，$h(y)$是$g(x)$的反函数，

且$\boldsymbol{\alpha}=g(a), \boldsymbol{\beta}=g(b)$.

y= g(x) ⟺ x=h(y)

h′(y) = 1/g′(x) =1/g′(h(y))

证：X在(a,b)上分布，则$Y=g(X)$在(α,β)上分布.

当$y\le\alpha$时，$F_Y(y)=P(Y\le y)=0$; (不可能事件)

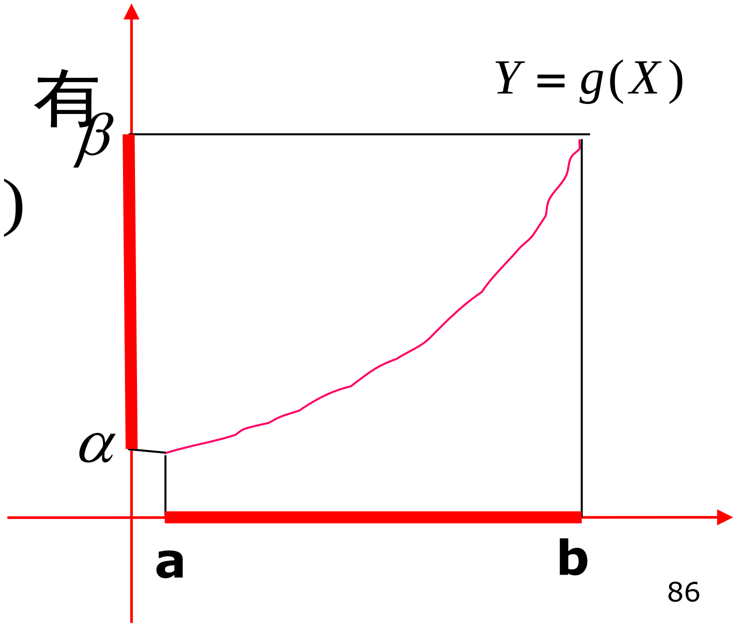
当$y\ge\beta$时，$F_Y(y)=P(Y\le y)=1$; (必然事件)

当$y\in(\alpha,\beta)$时，

$$F_Y(y)=P(Y\le y)=P(g(X)\le y)$$
$$=P(X\le h(y))=F_X(h(y)).$$

对y求导，用复合函数求导法则，有

$$f_Y(y)=\begin{cases} f_X(h(y))\cdot h'(y), & y\in(\alpha,\beta) \\ 0, & \text{其它.}\end{cases}$$



注：当$g(x)$严格下降时，

类似结论亦成立。

注记：

用概率微元法理解："概率微元不变"，即

X 落在小区间$(x, x+\Delta x)$上的概率

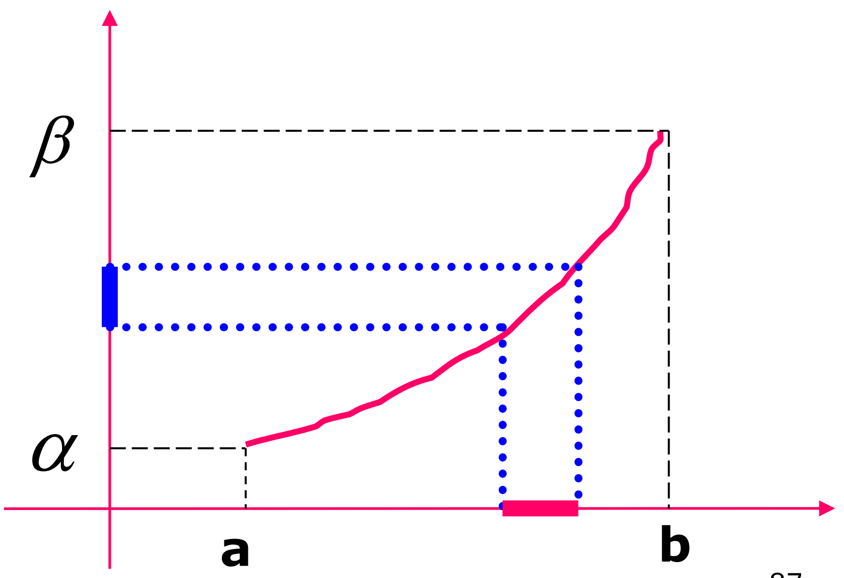
$=$ Y 落在小区间$(y, y+\Delta y)$上的概率

$$\Rightarrow P(x \le X \le x+dx) = P(y \le Y \le y+dy)$$

$$\Rightarrow f_X(x)dx = f_Y(y)dy$$

$$\Rightarrow f_Y(y) = f_X(x) \cdot \frac{dx}{dy}$$

$\frac{dx}{dy}$ 理解成：$\left|\frac{dx}{dy}\right|$.



例：正态分布的线性函数依然是正态分布

设$X \sim N(\mu,\sigma^2)$,求$Y = aX + b(a \neq 0)$的概率密度函数。

解：因$X \sim N(\mu,\sigma^2)$，其密度函数为：$f_X(x) = \dfrac{1}{\sqrt{2\pi}\ \sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R.$

函数$y = ax + b$为严格单调函数，

其反函数$x = h(y) = (y-b)/a$,导数$h'(y) = 1/a$. 由变换公式：

$$f_Y(y) = f_X(x)\cdot|h'(y)| = f_X((y-b)/a)\cdot 1/|a|$$

$\boxed{f_X(x)dx = f_Y(y)dy}$

$$= \frac{1}{\sqrt{2\pi}\ |a|\sigma} e^{-\frac{((y-b)/a-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\ |a|\sigma} e^{-\frac{(y-(a\mu+b))^2}{2a^2\sigma^2}}$$

$$\Rightarrow Y \sim N(a\mu + b, a^2\sigma^2).$$

正态变量的线性不变性

注：只要知道Y服从正态分布，其均值、方差可计算如下：

$$E(Y) = E(aX+b) = a\mu + b; \quad D(Y) = D(aX+b) = a^2\sigma^2.$$

例：对数正态分布 (Lognormal Distribution)

设$X \sim N(\mu,\sigma^2)$,定义$Y=e^X$，则$\ln Y = X \sim N(\mu,\sigma^2)$,

称Y服从对数正态分布。求Y的概率密度函数。

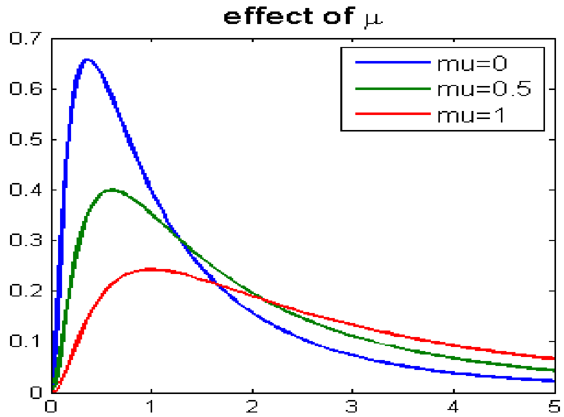
解：函数$y=e^x$严格单调上升，

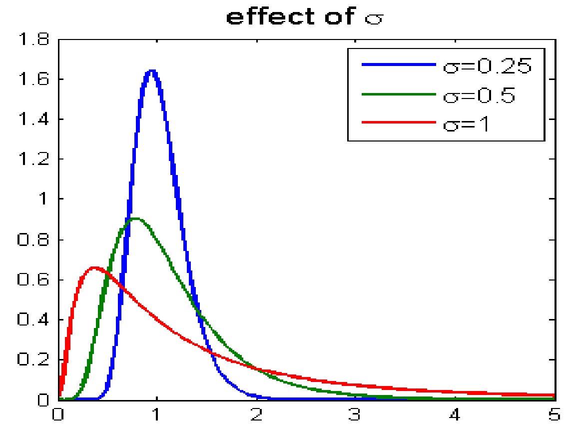
反函数$x=\ln y, y>0.$

由 $f_X(x)dx = f_Y(y)dy$

$$\Rightarrow f_Y(y) = f_X(x)\cdot\left|\frac{dx}{dy}\right|$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}\cdot\frac{1}{y},\quad y>0.$$





例：设T时刻股票价格

$$S_T = S_0 \exp((r-\boldsymbol{\sigma}^2/2)T + \boldsymbol{\sigma}\sqrt{T}\ Z), Z \sim N(0,1).$$

(r,σ **为参数)，**

则 $\ln S_T = \ln S_0 + (r-\boldsymbol{\sigma}^2/2)T + \boldsymbol{\sigma}\sqrt{T}\ Z.$

$\Rightarrow \ln S_T \sim N(\ln S_0 + (r-\boldsymbol{\sigma}^2/2)T, \boldsymbol{\sigma}^2 T).$

股票价格服从对数正态分布。

$$\mathbf{E}\left(S_T\right) = \int_{-\infty}^{+\infty} S_0 \exp((r-\frac{\boldsymbol{\sigma}^2}{2})T + \boldsymbol{\sigma}\sqrt{T}\ \mathbf{z})\frac{\mathbf{1}}{\sqrt{\mathbf{2}\boldsymbol{\pi}}}\mathbf{exp(-}\frac{\mathbf{z}^2}{2}\mathbf{)\,dz}$$

$$= S_0\, e^{rT}.$$

3. 变换法推广：定理

Many-to-one function

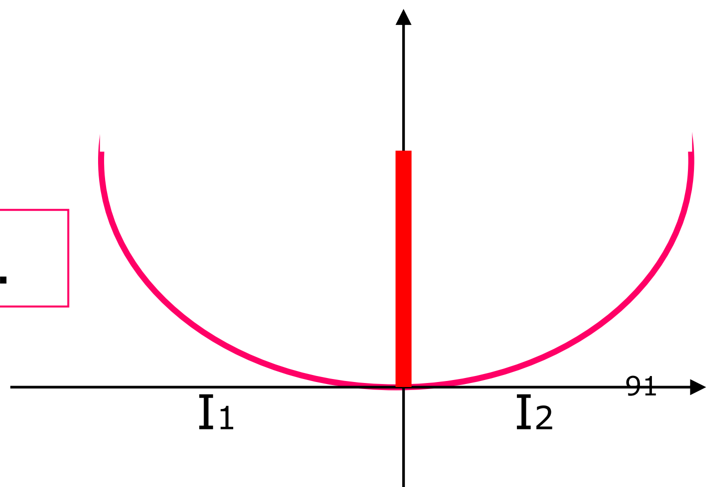
Y= g(x)

设连续型**RV** X具有密度$f_X(x)$. 在不相交的区间$I_1, I_2, \cdots$上逐段严格单调，在每一段上函数$y=g(x)$的值域均为D,反函数分别为$h_1(y)$，$h_2(y), \cdots$

则$Y := g(X)$为连续型RV,其概率密度函数为

$$f_Y(y)=\begin{cases}\sum_k f_X(h_k(y))\cdot|h_k'(y)|, y\in D;\\ 0, \qquad\qquad \text{其它.}\end{cases}$$

y=g(x) ⟺ x = h_k(y), k=1, 2,...



证明：

给定实数a，记$E_k(a)=\{x: x\in I_k, g(x)\le a\}$.

$$F_Y(a)=P(Y\le a)=P(g(X)\le a)=P(X\in\bigcup_k E_k(a))$$

$$=\sum_k P(X\in E_k(a))=\sum_k\int_{E_k(a)} f_X(x)dx$$

$$=\sum_k\int_{-\infty}^{a} f_X(h_k(y))|h_k'(y)|dy$$ 积分变换

$$=\int_{-\infty}^{a}\underline{\sum_k f_X(h_k(y))|h_k'(y)|}dy.$$

$$\Rightarrow q_Y(y)=f_X(h_k(y))|h_k'(y)|.$$

例： 设 $X \sim N(0,1)$，求 $Y = X^2$ 的概率密度函数。

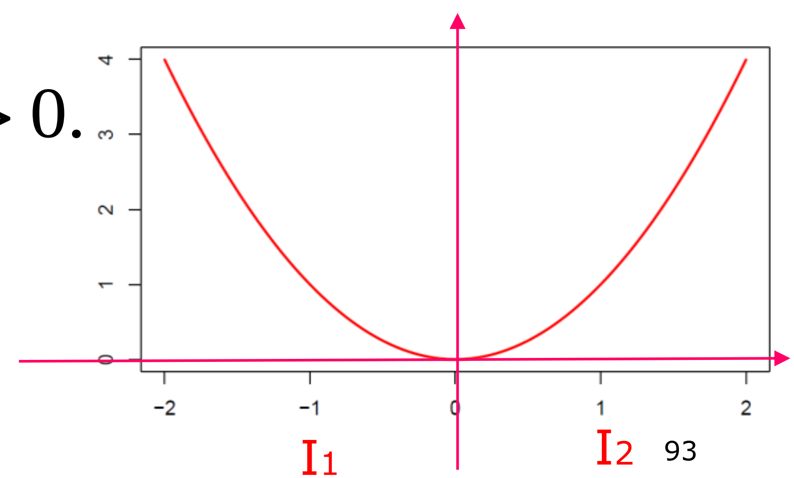
解：$y = x^2$ 分段严格单调：

在 $I_1 = (-\infty, 0)$ 上，反函数 $x = h_1(y) = -\sqrt{y}, y > 0$;

在 $I_2 = (0, +\infty)$ 上，反函数 $x = h_2(y) = \sqrt{y}, y > 0$;

$$\Rightarrow f_Y(y) = f_X(h_1(y)) \cdot \left| -\frac{1}{2\sqrt{y}} \right| + f_X(h_2(y)) \cdot \left| \frac{1}{2\sqrt{y}} \right|$$

$$= \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, y > 0.$$



与用直接法得到的结果相同

例:

设$X \sim f_X(x)=\begin{cases}2x/\pi^2, 0<x<\pi;\\ 0, 其它\end{cases}$ 求$Y=\sin X$的概率密度函数。

解：X在$(0,\pi)$内取值，$Y=\sin X$在$(0,1)$内取值。

函数$y=\sin x$分段严格单调:

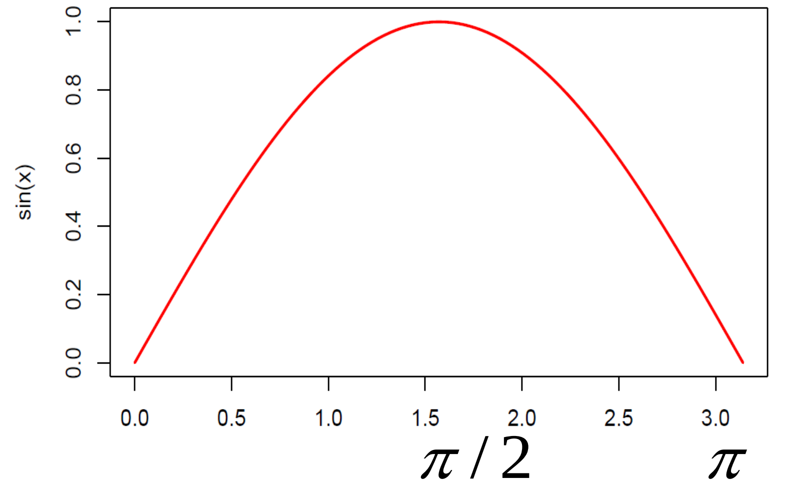
在$(0,\pi/2)$上，严格上升，反函数为$x=\arcsin y =: h_1(y)$;

在$(\pi/2,\pi)$上，严格下降，反函数为$x=\pi-\arcsin y =: h_2(y)$;

$$f_Y(y)=f_X(h_1(y))\cdot|h'_1(y)|+f_X(h_2(y))\cdot|h'_2(y)|$$

$$=\frac{2\arcsin y}{\pi^2\sqrt{1-y^2}}+\frac{2(\pi-\arcsin y)}{\pi^2\sqrt{1-y^2}}$$

$$=\frac{2}{\pi\sqrt{1-y^2}}, 0<y<1.$$



The End of Chapter 2