概率论与数理统计

何志坚

华南理工大学数学学院

hezhijian@scut.edu.cn

第四章：随机变量的数字特征

随机变量的自我介绍

随机变量X：本人随机变量X，连续型，期望为0.1，方差为1.2，偏度为-1.1，峰度为2.2....

随机变量Y：我是随机变量Y, 也是连续型，与随机变量X是好朋友，我们两的相关系数是0.99. 我觉得X是个很好的随机变量！

Dr. He: 好与坏关键要看是否“合适”。

目录

一维随机变量的数字特征

概率分布全面地描述了随机变量的统计规律性，由分布可以计算有关随机变量事件的概率。除此之外，由分布还可以计算随机变量的一些数字特征。这些特征从侧面描述了分布的特征。

- 期望：描述分布的“平均水平”
- 方差：描述分布的“波动大小”

环数	8	9	10
甲	0.1	0.8	0.1
乙	0.8	0.1	0.1
丙	0.2	0.6	0.2

如何比较甲、乙、丙三个选手的水平高低?

离散型随机变量数学期望的定义

定义：设离散随机变量X的分布列为

$$p_i = p(x_i) = P(X = x_i),\ i = 1, 2, \ldots.$$

如果级数$\sum_{i=1}^{\infty} |x_i|\, p(x_i) < \infty$[1], 则称$X$数学期望存在，并称

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$

为X的数学期望(Expectation)，简称X的期望。

环数	8	9	10	期望
甲	0.1	0.8	0.1	$8 \times 0.1 + 9 \times 0.8 + 10 \times 0.1 = 9.0$
乙	0.8	0.1	0.1	$8 \times 0.8 + 9 \times 0.1 + 10 \times 0.1 = 8.3$
丙	0.2	0.6	0.2	$8 \times 0.2 + 9 \times 0.6 + 10 \times 0.2 = 9.0$

[1]要求级数绝对收敛的目的在于使数学期望唯一。

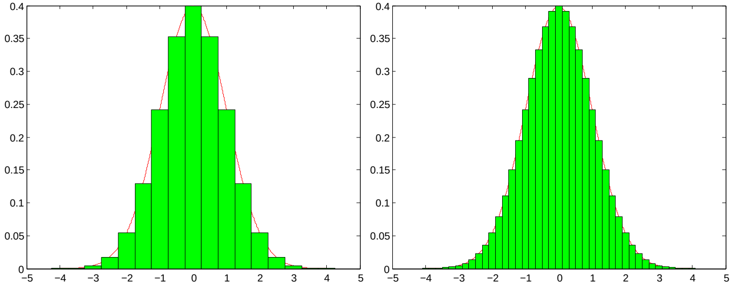
连续型随机变量数学期望的定义

定义： 设连续型随机变量X的密度函数为$p(x)$. 如果

$$\int_{-\infty}^{\infty} |x|\, p(x)\, \mathrm{d}x < \infty,$$

则称X**数学期望**存在，定义为

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x p(x)\, \mathrm{d}x.$$



随机变量函数的数学期望

定理：随机变量X的某一函数$g(X)$的数学期望为

$$\mathbb{E}[g(X)] = \begin{cases} \sum_{i=1}^{\infty} g(x_i)p(x_i) & \text{离散型} \\ \int_{-\infty}^{\infty} g(x)p(x)\,\mathrm{d}x & \text{连续型} \end{cases}$$

期望的性质：

- 如果c为常数，则$\mathbb{E}(c) = c$.
- 对任意常数a, 有$\mathbb{E}(aX) = a\mathbb{E}(X)$.
- 对于任意两个函数g_1, g_2, 有

$$\mathbb{E}[g_1(X) \pm g_2(X)] = \mathbb{E}[g_1(X)] \pm \mathbb{E}[g_2(X)].$$

- 如果$a \le X \le b$, 则$\mathbb{E}(X)$存在，且$a \le \mathbb{E}(X) \le b$.

随机变量的方差与标准差

定义： 设X为一随机变量。如果$\mathbb{E}(X^2) < \infty$, 则称偏差平方$(X - \mathbb{E}(X))^2$的数学期望$\mathbb{E}[(X - \mathbb{E}(X))^2]$为随机变量$X$的方差(Variance)，记为

$$\mathrm{Var}(X) = D(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \begin{cases} \sum_{i=1}(x_i - \mathbb{E}(X))^2 p(x_i) & \text{离散型} \\ \int_{-\infty}^{\infty}(x - \mathbb{E}(X))^2 p(x)\,\mathrm{d}x & \text{连续型}. \end{cases}$$

称方差的正平方根$\sqrt{\mathrm{Var}(X)}$为标准差，记为$\sigma(X)$或σ_X.

注： 方差与标准差用来描述随机变量取值的集中与分散程度，如果他们越小则说明随机变量的取值越集中；反之越发散。

环数	8	9	10	期望	方差
甲	0.1	0.8	0.1	9.0	$1^2 \times 0.1 + 1^2 \times 0.1 = 0.20$
乙	0.8	0.1	0.1	8.3	$0.3^2 \times 0.8 + 0.7^2 \times 0.1 + 1.7^2 \times 0.1 = 0.41$
丙	0.2	0.6	0.2	9.0	$1^2 \times 0.2 + 1^2 \times 0.2 = 0.40$

方差的性质

方差的等价表达形式

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

- 如果c为常数，则$\mathrm{Var}(c) = 0$, $\mathrm{Var}(X + c) = \mathrm{Var}(X)$.
- 对任意常数a, 有$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$.
- $\mathrm{Var}(X) = 0$的等价条件是存在常数c使得$P(X = c) = 1$.

二项分布的期望与方差

定理： 如果$X \sim B(n,p)$, 则$\mathbb{E}[X] = np$, $\mathrm{Var}[X] = np(1-p)$.

证： 令$q = 1-p$.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=0}^{n} kB(k;n,p) = \sum_{k=1}^{n} kC_n^k p^k q^{n-k} \\ &= np\sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{n-k} \\ &= np\sum_{k=0}^{n-1} C_{n-1}^{k} p^{k} q^{n-1-k} = np\end{aligned}$$

其中用到$kC_n^k = nC_{n-1}^{k-1}$

二项分布的期望与方差

定理: 如果$X \sim B(n, p)$, 则$\mathbb{E}[X] = np$, $\mathrm{Var}[X] = np(1-p)$.

证:

$$\begin{aligned}
\mathbb{E}[X^2] &= np\sum_{k=1}^{n} kC_{n-1}^{k-1}p^{k-1}q^{n-k} \\
&= np\sum_{k=0}^{n-1}(k+1)C_{n-1}^{k}p^{k}q^{n-1-k} \\
&= np\left(\sum_{k=1}^{n-1} kC_{n-1}^{k}p^{k}q^{n-1-k} + 1\right) \\
&= np((n-1)p+1) = (np)^2 + npq.
\end{aligned}$$

其中用到$\sum_{i=1}^{n} iC_n^i p^{i-1}q^{n-i} = n(p+q)^{n-1}$. 所以

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = npq = np(1-p).$$

泊松分布的期望与方差

定理： 如果$X \sim Pois(\lambda)$, 则$\mathbb{E}[X] = \mathrm{Var}[X] = \lambda$.

证：

$$\mathbb{E}[X] = e^{-\lambda}\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^k}{k!}$$

$$= \lambda e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

$$\mathbb{E}[X^2] = e^{-\lambda}\sum_{k=0}^{\infty} k^2\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty}(k+1)\frac{\lambda^k}{k!}$$

$$= \lambda e^{-\lambda}\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!} + \lambda = \lambda^2 + \lambda$$

$$\mathrm{Var}[x] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \lambda$$

几何分布的期望与方差

定理： 如果$X \sim Geo(p)$, 则$\mathbb{E}[X] = \frac{1}{p}$，$\mathrm{Var}[X] = \frac{1-p}{p^2}$.

证： 设$q = 1 - p$.

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} kq^{k-1}p = p\frac{1}{(1-q)^2} = \frac{1}{p}.$$

$$\begin{aligned}
\mathbb{E}[X^2] &= p\sum_{k=1}^{\infty} k^2 q^{k-1} = p\sum_{k=1}^{\infty}[k(k-1)+k]q^{k-1} \\
&= pq\sum_{k=2}^{\infty} k(k-1)q^{k-2} + \frac{1}{p} \\
&= pq\frac{2}{(1-q)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}
\end{aligned}$$

$$\mathrm{Var}[x] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}$$

均匀分布的期望与方差

定理： 如果$X \sim \mathbb{U}[a,b]$, 则$\mathbb{E}[X] = (b-a)/2$, $\mathrm{Var}[X] = (b-a)^2/12$.

证：

$$\mathbb{E}[X] = \int_a^b \frac{x}{b-a}\,\mathrm{d}x = \frac{a+b}{2}$$

$$\mathbb{E}[X^2] = \int_a^b \frac{x^2}{b-a}\,\mathrm{d}x = \frac{a^2+ab+b^2}{3}$$

$$\mathrm{Var}[x] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

指数分布的期望与方差

定理：如果$X \sim Exp(\lambda)$, 则$\mathbb{E}[X] = 1/\lambda$, $\mathrm{Var}[X] = (b-a)^2/12$.

证：

$$\begin{aligned}\mathbb{E}[X] &= \int_0^{\infty} x\lambda e^{-\lambda x}\,\mathrm{d}x = -\int_0^{\infty} x\,\mathrm{d}e^{-\lambda x} \\ &= -xe^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda} \ \text{(分部积分)}\end{aligned}$$

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^{\infty} x^2\lambda e^{-\lambda x}\,\mathrm{d}x \\ &= -x^2 e^{-\lambda x}\Big|_0^{\infty} + 2\int_0^{\infty} xe^{-\lambda x}\,\mathrm{d}x \ \text{(分部积分)} \\ &= \frac{2}{\lambda^2}\end{aligned}$$

$$\mathrm{Var}[x] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1}{\lambda^2}$$

正态分布的期望与方差

定理： 如果$X \sim N(\mu, \sigma^2)$, 则$\mathbb{E}[X] = \mu$, $\mathrm{Var}[X] = \sigma^2$.

证：

$$\mathbb{E}[X - \mu] = \int_{-\infty}^{\infty} (x - \mu) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, \mathrm{d}x = 0.$$

所以$\mathbb{E}[X] = \mu$.

$$\begin{aligned}
\mathrm{Var}[x] &= \int_{-\infty}^{\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, \mathrm{d}x \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} \, \mathrm{d}x \\
&= -\frac{\sigma^2}{\sqrt{2\pi}} x e^{-x^2/2} \Big|_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} \, \mathrm{d}x \ \text{(分部积分)} \\
&= \sigma^2.
\end{aligned}$$

柯西分布的期望与方差

定理： 柯西分布的期望与方差不存在。

证： 设X服从柯西分布，其密度函数为

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$$

$$\begin{aligned}
\mathbb{E}[|X|] &= \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)} \mathrm{d}x = \frac{2}{\pi}\int_0^{\infty} \frac{x}{1+x^2} \mathrm{d}x \\
&= \frac{1}{\pi}\int_0^{\infty} \frac{1}{1+y} \mathrm{d}y \\
&= \frac{1}{\pi}\int_1^{\infty} \frac{1}{y} \mathrm{d}y = \infty
\end{aligned}$$

总结

分布名称	记号	期望	方差
二项分布	$B(n,p)$	np	npq
泊松分布	$Pois(\lambda)$	λ	λ
几何分布	$Geo(p)$	$1/p$	q/p^2
均匀分布	$U[a,b]$	$(a+b)/2$	$(b-a)^2/12$
指数分布	$Exp(\lambda)$	$1/\lambda$	$1/\lambda^2$
正态分布	$N(\mu,\sigma^2)$	μ	σ^2
柯西分布		不存在	不存在

作业：计算上述前六种分布的期望与方差。**5月11日(周五)交！**

随机变量的矩

定义： 设X为随机变量，c为常数，k为正整数，如果$\mathbb{E}[|X-c|^k]<\infty$, 则称

$$\mathbb{E}[(X-c)^k]$$

为X关于点c的k阶矩。特别地，

(1) 当$c=0$，称$\mathbb{E}[X^k]$为X的k阶原点矩；

(2) 当$c=\mathbb{E}[X]$，称$\mathbb{E}[(X-\mathbb{E}[X])^k]$为$X$的$k$阶中心矩。

注： 期望是一阶原点矩；方差是二阶中心矩。

标准正态分布的矩

定理： 如果$X \sim N(0,1)$, 对任意的正整数k，有

$$\mathbb{E}[X^k] = \begin{cases} 0 & k\text{为奇数} \\ (k-1)!! & k\text{为偶数} \end{cases}$$

证： 当k为奇数时，$\mathbb{E}[X^k] = \mathbb{E}[(-X)^k] = -\mathbb{E}[X^k]$. 所以$\mathbb{E}[X^k] = 0$. 当$k$为偶数时，

$$\begin{aligned} \mathbb{E}[X^k] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2}\,\mathrm{d}x = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{k-1}\,\mathrm{d}e^{-x^2/2} \\ &= -\frac{1}{\sqrt{2\pi}} x^{k-1} e^{-x^2/2}\Big|_{-\infty}^{\infty} + \frac{k-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{k-2} e^{-x^2/2}\,\mathrm{d}x \\ &= (k-1)\mathbb{E}[X^{k-2}] \end{aligned}$$

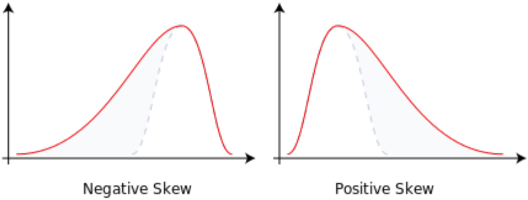
所以$\mathbb{E}[X^k] = (k-1)!!\mathbb{E}[X^0] = (k-1)!!$.

随机变量的偏度

定义： 设X为随机变量，如果$\mathbb{E}[|X|^3] < \infty$且$\mathrm{Var}[x] > 0$, 则称

$$\beta_s := \frac{\mathbb{E}[(X - \mathbb{E}[X])^3]}{(\mathrm{Var}[X])^{3/2}} = \mathbb{E}\left[\left(\frac{X - \mathbb{E}[X]}{\sigma_X}\right)^3\right]$$

为X的偏度(skewness). 如果$\beta_s < 0$时，称该分布左(负)偏；如果$\beta_s > 0$时，称该分布右(正)偏。



注： 偏度β_s是刻画分布偏离对称性程度的一个特征数。定义中的分母的作用是为了消除量纲的影响。

随机变量的峰度

定义：设X为随机变量，如果$\mathbb{E}[X^4] < \infty$且$\mathrm{Var}[x] > 0$, 则称下式为X的**峰度(kurtosis)**：

$$\beta_k = \frac{\mathbb{E}[(X - \mathbb{E}[X])^4]}{(\mathrm{Var}[X])^2} = \mathbb{E}\left[\left(\frac{X - \mathbb{E}[X]}{\sigma_X}\right)^4\right].$$

注：偏度β_k是刻画分布尖峭程度和（或）尾部粗细的一个特征数。通常与标准正态的峰度3比较。如果$\beta_k > 3$, 说明标准化后的分布比标准正态分布更尖峭和（或）尾部更粗；如果$\beta_k < 3$, 说明标准化后的分布比标准正态分布更平坦和（或）尾部更细。如果$\beta_s \approx 0, \beta_k \approx 3$, 常认为该分布为近似正态分布。

不同分布的峰度

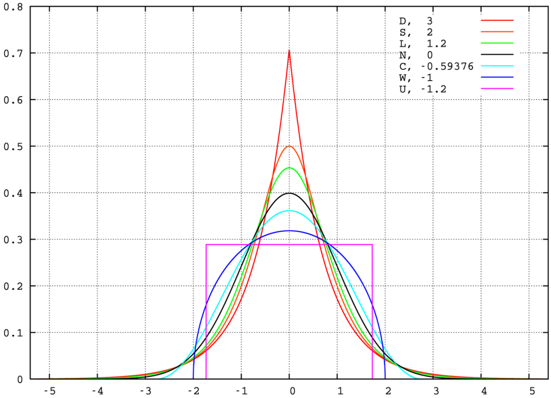


图: 图中为7种期望为0且方差为1的不同分布示意图，其他U表示均匀分布，N表示标准正态分布。右上角的数值表示$\beta_k - 3$的值

分位数

定义： 设随机变量X的分布函数为$F(x)$, 密度函数为$p(x)$. 对任意$p \in (0,1)$, 称满足条件

$$F(x_p) = \int_{-\infty}^{x_p} p(x)\,\mathrm{d}x = p$$

的x_p为此分布的p分位数(quantile). 当$p = 0.5$时，x_p称为中位数。

例： 标准正态分布的分位数：

p	0.01	0.05	0.1	0.5	0.9	0.95	0.99
x_p	-2.33	-1.64	-1.28	0	1.28	1.64	2.33

注意到$x_{1-p} = -x_p$

目录

二维随机向量函数的方差与期望

推广： 二维随机向量(X, Y)的某一函数$g(X, Y)$的数学期望为

$$\mathbb{E}[g(X,Y)] = \begin{cases} \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j)P(X = x_i, Y = y_j) & \text{离散型} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p(x,y)\,\mathrm{d}x\,\mathrm{d}y & \text{连续型} \end{cases}$$

类似地，

$$\mathrm{Var}[g(X,Y)] = \mathbb{E}[(g(X,Y) - \mathbb{E}[g(X,Y)])^2] = \mathbb{E}[g(X,Y)^2] - (\mathbb{E}[g(X,Y)])^2$$

性质：

- 线性性：$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.
- 如果X, Y独立，$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, $\mathrm{Var}[aX + bY] = a^2\mathrm{Var}[X] + b^2\mathrm{Var}[Y]$.
- 上述定义以及性质可以推广到n维随机向量。

二维随机向量的协方差

定义： 设(X, Y)为二维随机向量，且$\mathrm{Var}[X] > 0$, $\mathrm{Var}[Y] > 0$. 其协方差(covariance)定义为

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

相关系数定义为

$$r(X, Y) = \rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}[X]}\sqrt{\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

如果$r(X, Y) = 0$, 称X, Y不相关。

协方差的性质(I)

- 等价表达形式:
$$\mathrm{Cov}(X,Y)=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]$$
- 对称性:
$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$$
- 双线性性:
$$\mathrm{Cov}(aX+bY,Z)=a\mathrm{Cov}(X,Z)+b\mathrm{Cov}(Y,Z)$$
$$\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$$
-
$$\mathrm{Var}[X]=\mathrm{Cov}(X,X)$$
$$\mathrm{Var}(X+Y)=\mathrm{Var}[X]+\mathrm{Var}[Y]+2\mathrm{Cov}(X,Y)$$
- 如果X,Y独立，$\mathrm{Var}(X,Y)=0$.

施瓦茨(Schwarz)不等式

定理： 设(X, Y)为二维随机向量，则有

$$(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2]\mathbb{E}[Y^2]$$

证： 令

$$g(t) = \mathbb{E}[(tX + Y)^2] = \mathbb{E}[X^2]t^2 + 2\mathbb{E}[XY]t + \mathbb{E}[Y^2] \geq 0.$$

如果$\mathbb{E}[X^2] = 0$, 则意味着$P(X = 0) = 1$. 所以$\mathbb{E}[XY] = 0$.
如果$\mathbb{E}[X^2] > 0$, 则$g(t)$的判别式不大于0，即

$$(2\mathbb{E}[XY])^2 - 4\mathbb{E}[X^2]\mathbb{E}[Y^2] \leq 0.$$

推论： 设(X, Y)为二维随机向量，则有

$$[\mathrm{Cov}(X, Y)]^2 \leq \mathrm{Var}[X]\mathrm{Var}[Y].$$

相关系数的性质

定理： 设(X, Y)为二维随机向量，且$\mathrm{Var}[X] > 0$, $\mathrm{Var}[Y] > 0$.

- 如果X, Y独立，则$r(X, Y) = 0$, 即X, Y不相关。
- $$|r(X, Y)| \le 1$$
- $r(X, Y) = 1$的充要条件是存在常数$a > 0$和b, 使得
$$P(Y = aX + b) = 1$$
$r(X, Y) = -1$的充要条件是存在常数$a < 0$和b, 使得
$$P(Y = aX + b) = 1$$

证： 注意到$\mathrm{Var}(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}) = 2[1 \pm r(X, Y)]$

注： 相关系数反映X与Y线性相关的程度。若$r = 1$，称X, Y正相关，若$r = -1$，称X, Y负相关。

二维正态分布的相关系数

例： 设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. 则

$$r(X,Y)=\rho.$$

证明：

$$\mathrm{Cov}(X,Y)=\mathbb{E}[(X-\mu_1)(Y-\mu_2)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$$=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xye^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}\,\mathrm{d}x\,\mathrm{d}y=\rho\sigma_1\sigma_2.$$

$$r(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sigma_1\sigma_2}=\rho$$

注： 由此可得，对于二维正态分布，独立性与不相关性是等价的。但一般情况下，这种等价关系不存在（独立可以得到不相关，反之不成立）

反例

设(X,Y)的联合密度函数为

$$f(x,y)=\frac{3}{8\pi\sqrt{2}}\left\{\exp\left(-\frac{9}{16}(x^2-\frac{2}{3}xy+y^2)\right)+\exp\left(-\frac{9}{16}(x^2+\frac{2}{3}xy+y^2)\right)\right\}$$

求$r(X,Y)$.

解： 易知，$X\sim N(0,1), Y\sim N(0,1)$. 容易发现$\mathbb{E}[XY]=0$.
又$\mathbb{E}[X]=\mathbb{E}[Y]=0$. 由此可得$r(X,Y)=0$.

注： 虽然$r(X,Y)=0$，但X,Y不独立。这是因为$f(x,y)\neq f_X(x)f_Y(y)$.

例

例：已知X, Y相互独立，均服从正态分布$N(0, \sigma^2)$，$U = aX + bY$，$V = aX - bY$，其中a, b是常数。

- 求U, V的相关系数
- U, V是否相关，是否独立？
- 当U, V独立时，求它们的联合密度函数。

解：$\mathbb{E}[UV] = \mathbb{E}[(aX + bY)(aX - bY)] = a^2\mathbb{E}[X^2] - b^2\mathbb{E}[Y^2] = (a^2 - b^2)\sigma^2$. 由于$\mathbb{E}[U] = \mathbb{E}[V] = 0$, 所以$\mathrm{Cov}(X, Y) = (a^2 - b^2)\sigma^2$.

$$\mathrm{Var}(U) = \mathrm{Var}(V) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) = (a^2 + b^2)\sigma^2.$$

$$r(X, Y) = \frac{a^2 - b^2}{a^2 + b^2}$$

协方差矩阵

定义： 设随机向量$\mathbf{X}=(X_1,\ldots,X_n)^\top$的每个分量都有有限方差，定义协方差矩阵如下

$$\mathrm{Var}(\mathbf{X})=\begin{pmatrix} \mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n) \end{pmatrix}$$

相关系数矩阵定义如下：

$$r(\mathbf{X})=\begin{pmatrix} 1 & r(X_1,X_2) & \cdots & r(X_1,X_n) \\ r(X_2,X_1) & 1 & \cdots & r(X_2,X_n) \\ \vdots & \vdots & \vdots & \vdots \\ r(X_n,X_1) & r(X_n,X_2) & \cdots & 1 \end{pmatrix}$$

条件数学期望

定义： 设(X,Y)为二维随机向量，有有限的数学期望。在$\{Y=y\}$发生的条件下，X的条件数学期望（简称条件期望），就是在条件分布$X|Y=y$ 下求条件期望。

$$\mathbb{E}[X|Y=y]=\begin{cases}\sum_{i=1}^{\infty} x_i P(X=x_i|Y=y) & \text{离散型}\\ \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,\mathrm{d}x & \text{连续型}\end{cases}$$

同样方式定义$\mathbb{E}[Y|X=x]$.

注1： $\mathbb{E}[X|Y=y]$可以看成关于y的函数，记为$g(y)=\mathbb{E}[X|Y=y]$.
则$g(Y)=\mathbb{E}[X|Y]$为随机变量Y的函数。同样$\mathbb{E}[Y|X]$为随机变量X的函数。

注2： 如果X,Y独立$\mathbb{E}[X|Y]=\mathbb{E}[X]$，$\mathbb{E}[Y|X]=\mathbb{E}[Y]$.

注3： 条件期望满足前面提到期望的所有性质，比如线性性。

全期望公式

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$$

证明： 考虑离散的情形。

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \sum_{j=1}^{\infty} \mathbb{E}[X|Y=y_j]P(Y=y_j) \\
&= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} x_i P(X=x_i|Y=y_j)P(Y=y_j) \\
&= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} x_i P(X=x_i, Y=y_j) \\
&= \sum_{i=1}^{\infty} x_i \sum_{j=1}^{\infty} P(X=x_i, Y=y_j) \\
&= \sum_{i=1}^{\infty} x_i P(X=x_i) = \mathbb{E}[X]
\end{aligned}$$

全期望公式

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$$

证明： 考虑连续的情形。

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y\right)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \mathbb{E}[X]
\end{aligned}$$

例1：二维正态向量的条件期望

例： 若$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，求$\mathbb{E}[X|Y], \mathbb{E}[Y|X]$.

解： 前一章已经证明：

$$X|Y = y \sim N(\mu_1 + \rho(y - \mu_2)\sigma_1/\sigma_2, \sigma_1^2(1 - \rho^2)),$$

$$Y|X = x \sim N(\mu_2 + \rho(x - \mu_1)\sigma_2/\sigma_1, \sigma_2^2(1 - \rho^2)).$$

所以，$\mathbb{E}[X|Y] = \mu_1 + \rho(Y - \mu_2)\sigma_1/\sigma_2$. $\mathbb{E}[Y|X] = \mu_2 + \rho(X - \mu_1)\sigma_2/\sigma_1$.

例2：随机个随机变量的和

例： 设$X_1, X_2, \cdots$为独立同分布随机变量序列，期望为μ_1，N为非负整数值随机变量，期望为μ_2，且与$X_1, X_2, \cdots$独立，求$\mathbb{E}[\sum_{k=1}^{N} X_k]$.

解： 由全期望公式可得，

$$
\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N} X_k\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{N} X_k \middle| N\right]\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{N} X_k \middle| N = n\right] P(N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{n} X_k\right] P(N = n) \\
&= \sum_{n=0}^{\infty} \mu_1 n P(N = n) = \mu_1 \sum_{n=0}^{\infty} n P(N = n) = \mu_1 \mu_2 .
\end{aligned}
$$

条件方差

定义： 条件方差定义：

$$\mathrm{Var}[X|Y] = \mathbb{E}[(\mathbb{E}[X|Y] - \mathbb{E}[X])^2] = \mathbb{E}[(\mathbb{E}[X|Y])^2] - (\mathbb{E}[X])^2$$

性质：

$$\mathrm{Var}[X] = \mathrm{Var}[\mathbb{E}[X|Y]] + \mathbb{E}[\mathrm{Var}[X|Y]]$$

作业

第四章习题：2, 8, 13, 14, 16, 21, 22

5月18日（周五）交！