第十六讲：数学期望

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

有甲乙两名射手，其射击技术可用下表表出

甲射手

击中环数	8	9	10
概　率	0.3	0.1	0.6

乙射手

击中环数	8	9	10
概　率	0.2	0.5	0.3

试问哪一个射手技术较好?

- 显然这个问题的答案不是一眼就可以看出来的;
- 这也表明分布列虽然完整的描述了随机变量，但却不够集中的反应其变化情况;
- 有必要找一些量来更集中，更概括的描述随机变量;
- 所要找的量多是某种平均值.

平均值与加权平均值

- 最常见的平均值求法:$\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n}$;
- 没有考虑每个数据相对重要性，比如一小学生考试成绩为：语文 95 分，数学 85 分，常识 60 分，若按上述计算其平均成绩为 $\bar{x} = 80$;
- 上述计算没有考虑到三个科目的相对重要性，不太能反映学生的真正成绩：如果在这个年级中，每周有语文 10 节课，数学 8 节课，常识 2 节课，则利用下述的平均计算其平均成绩似科更合理一些:

$$\bar{x}_w = 95 * \frac{10}{20} + 85 * \frac{8}{20} + 60 * \frac{2}{20} = 87.5$$

- 加权平均：给定权 w_i 满足 $\sum_{i=1}^{n} w_i = 1$, 则

$$\bar{x}_w = \sum_{i=1}^{n} w_i x_i$$

称为加权平均值.

有甲乙两名射手，其射击技术可用下表表出

甲射手

击中环数	8	9	10
概　　率	0.3	0.1	0.6

乙射手

击中环数	8	9	10
概　　率	0.2	0.5	0.3

考虑以概率为权重的加权平均:

$$\bar{x}_{\text{甲}} = 8 * 0.3 + 9 * 0.1 + 10 * 0.6 = 9.3$$
$$\bar{x}_{\text{乙}} = 8 * 0.2 + 9 * 0.5 + 10 * 0.3 = 9.1$$

则平均起来，甲每枪射中 9.3 环，乙每枪射中 9.1 环，故甲的技术更好一些.

定义 1.1 设离散型随机变量的分布列为

$$\begin{pmatrix} x_1, & x_2, & \cdots, & x_k, & \cdots \\ p_1, & p_2, & \cdots, & p_k, & \cdots \end{pmatrix}$$

如果

$$\sum_{i=1}^{\infty} |x_i| p_i < \infty$$

则称

$$E(X) = \sum_{i=1}^{\infty} x_i P(X = x_i) = \sum_{i=1}^{\infty} x_i p_i$$

为随机变量 X 的数学期望，或称该分布的数学期望，简称期望或均值。若级数 $\sum_{i=1}^{\infty} |x_i| p_i$ 不收敛，则称 X 的期望不存在.

从离散到连续

- 假设连续型随机变量 X 的分布密度为 $p(x)$;
- 考虑随机变量 X 的如下近似：记 $A_i := \{\omega : X(\omega) \in (x_i, x_{i+1}]\}$

$$\tilde{X} := \sum_i x_i I_{A_i}(\omega)$$

其中 $I_{A_i}(\omega) = \begin{cases} 1, & \omega \in A_i \\ 0, & \omega \notin A_i \end{cases}.$
显然 $\tilde{X}$ 为离散型随机变量，且

$$P(\tilde{X} = x_i) = \int_{x_i}^{x_{i+1}} p(x)dx \approx p(x_i)(x_{i+1} - x_i);$$

- 故 $\tilde{X}$ 的期望为

$$E(\tilde{X}) \approx \sum_i x_i p(x_i)(x_{i+1} - x_i) \to \int xp(x)dx$$

连续型随机变量数学期望的定义

定义 1.2 设连续随机变量 X 的密度函数为 $p(x)$. 如果

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

则称

$$E(X) := \int_{-\infty}^{\infty} xp(x)dx$$

为随机变量 X 的数学期望，简称期望或均值。若 $\int_{-\infty}^{\infty} |x|p(x)dx$ 不收敛，则称 X 的期望不存在.

- 若随机变量 X 的分布函数为 $F(x)$;
- 类似于连续情形，考虑随机变量 X 的如下近似:

$$\tilde{X} := \sum_i x_i I_{A_i}(\omega)$$

其中 $A_i := \{\omega : X(\omega) \in (x_i, x_{i+1}]\}, I_{A_i}(\omega) = \begin{cases} 1, & \omega \in A_i \\ 0, & \omega \notin A_i \end{cases}.$
显然 $\tilde{X}$ 为离散型随机变量，且

$$P(\tilde{X} = x_i) = P(X \in (x_i, x_{i+1}]) = F(x_{i+1}) - F(x_i);$$

- 故 $\tilde{X}$ 的期望为

$$E(\tilde{X}) = \sum_i x_i \left[F(x_{i+1}) - F(x_i)\right] \to \int x dF(x)$$ 斯蒂尔切斯积分

定义 1.3 假设随机变量 X 的分布函数为 $F(x)$, 则定义

$$E(X)=\int_{-\infty}^{+\infty} x dF(x)$$

为随机变量 X 的数学期望。这里我们要求上述积分绝对收敛，否则称数学期望不存在.

斯蒂尔切斯 (Stieltjes) 积分的性质

$$I=\int_{-\infty}^{+\infty} g(x)dF(x)$$

- 当 $F(x)$ 为右连续的阶梯函数，在 $x_i(i=1,2,\cdots,)$ 具有跃度 p_i 时，上面的积分化为

$$I=\sum_i g(x_i)\Big(F(x_i)-F(x_{i-1})\Big)=\sum_i g(x_i)p_i$$

- 当 $F(x)$ 存在导数 $F'(x)=p(x)$ 时，上述积分化为普通积分

$$I=\int_{-\infty}^{+\infty} g(x)p(x)dx$$

- 线性性质

$$\int_{-\infty}^{+\infty}(ag_1(x)+bg_2(x)dF(x)=a\int_{-\infty}^{+\infty} g_1(x)dF(x)+b\int_{-\infty}^{+\infty} g_2(x)dF(x)$$

- 若 $g(x)\geq 0$, $F(x)$ 单调不减，$b>a$, 则 $\int_a^b g(x)dF(x)\geq 0$

- 在引入上述相关的积分后，

$$
\begin{aligned}
F(x) &= P(X \le x) = \int_{(-\infty,x]} dF(x) \\
P(X \in B) &= \int_B dF(x)
\end{aligned}
$$

- 数学期望的另一计算公式：关于概率测度的积分

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xdF(x) \quad \text{注意到} F(x) = P(X(\omega) \le x) := P(X^{-1}(x)) \\
&\xlongequal{x=X(\omega)} \int_{\Omega} X(\omega)dP(X^{-1}(X(\omega))) = \int_{\Omega} X(\omega)dP(\omega)
\end{aligned}
$$

第十七讲：数学期望 (续) 与方差

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

- 若 $X(\omega)=\sum_{i=1}^{n} a_i I_{A_i}(\omega), A_1,\cdots,A_n\in\mathcal{F}$, 即 X 为简单随机变量时, 定义

$$\int_\Omega X(\omega)P(d\omega) := \sum_{i=1}^{n} a_i P(A_i)$$

- 若 $X(\omega)$ 为非负随机变量，则由之前随机变量的结构易知:

$$\text{存在简单随机变量序列} X_n(\omega)\uparrow X(\omega)$$

故可定义

$$\int_\Omega X(\omega)P(d\omega) := \lim_{n\to\infty}\int_\Omega X_n(\omega)P(d\omega)$$

- 若 $X(\omega)$ 为任一随机变量，则由于 $X(\omega)=X^+(\omega)-X^-(\omega)$, 故定义

$$\int_\Omega X(\omega)P(d\omega) := \int_\Omega X^+(\omega)P(d\omega)-\int_\Omega X^-(\omega)P(d\omega)$$

定义 1.4 称随机变量 $X(\omega)$ 关于概率测度 P 的积分存在，如果

$$\int_\Omega X^+(\omega)P(d\omega) \text{ 与 } \int_\Omega X^-(\omega)P(d\omega) \text{不同时为} \infty$$

称随机变量 $X(\omega)$ 关于概率测度 P 可积，如果 $\int_\Omega X(\omega)P(d\omega) < \infty$, 即

$$\int_\Omega X^+(\omega)P(d\omega) < \infty \text{ 且 } \int_\Omega X^-(\omega)P(d\omega) < \infty.$$

定义 1.5 若 $X(\omega)$ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的可积随机变量，则称

$$E(X) := \int_\Omega X(\omega)P(d\omega)$$

为随机变量 $X(\omega)$ 的数学期望.

随机变量期望的性质

1. 若 c 为常数，则 $E(c) = c$;
2. 若 $X \geq 0$, 则 $E(X) \geq 0$
3. 对任意的常数 a, 有 $E(aX) = aE(X)$;
4. 对任意两个函数 $g_1(x), g_2(x)$ 有

$$E(g_1(X) \pm g_2(X)) = E(g_1(X)) \pm E(g_2(X))$$

5. 对任意两个随机变量 X, Y, 有 $E(X + Y) = E(X) + E(Y)$;
6. 一般的，对任意两个随机变量 X, Y 及任给两个常数 a, b, 有

$$E(aX + bY) = aE(X) + bE(Y);$$

7. 更一般的有

$$E(\sum_{i=1}^{n} a_i X_i + b) = \sum_{i=1}^{n} a_i EX_i + b$$

积分变换定理

定理 1.1 设 X 为概率空间 $(\Omega, \mathcal{F}, P)$ 上的随机变量，$\mathbf{F}$ 为随机变量 X 的分布，则对任意的 Borel 函数 $f(x)$ 有

$$\int_{\Omega} f(X(\omega))P(d\omega) = \int_{\mathbb{R}} f(x)\mathbf{F}(dx) \left(= \int_{\mathbb{R}} f(x)dF(x) \right).$$

特别的，若 $f(x) = x$, 则有

$$E(X) = \int_{\Omega} X(\omega)P(d\omega) = \int_{\mathbb{R}} x\mathbf{F}(dx) = \int_{\mathbb{R}} xdF(x)$$

定理 1.2 设 X, Y 为概率空间 $(\Omega, \mathcal{F}, P)$ 上两个相互独立的随机变量，且 X, Y 均可积，则乘积 XY 也可积，且

$$E(XY) = E(X)E(Y)$$

定理 1.3 概率空间 $(\Omega, \mathcal{F}, P)$ 上两个随机变量 X, Y 相互独立的充要条件是，对于使得 $f(X)$ 与 $g(Y)$ 均可积的任何 Borel 函数 f, g 均有

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)].$$

定理 1.4 若随机变量 X 的分布函数为 $F(x)$, 则 X 的某一函数 $g(X)$ 的期望为

$$
\begin{aligned}
E[g(X)] &= \int_{-\infty}^{+\infty} g(x)dF(x) \\
&= \begin{cases} \sum_i g(x_i)[F(x_i) - F(x_{i-1})] = \sum_i g(x_i)P(X = x_i), \\ \\ \int_{-\infty}^{+\infty} g(x)p(x)dx. \end{cases}
\end{aligned}
$$

定理 1.5 若随机向量 $(X_1,\cdots,X_n)$ 的分布函数为 $F(x_1,\cdots,x_n)$，$g(x_1,\cdots,x_n)$ 为 n 元 Borel 函数，则

$$E[g(X_1,\cdots,X_n)] \quad = \quad \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n).$$

特别的，

$$E(X_i)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} x_i dF(x_1,\cdots,x_n)=\int_{-\infty}^{+\infty} x_i dF_i(x_i)$$

其中 $F_i(x_i)$ 为 X_i 的分布函数. 更进一步，如果 $(X_1,\cdots,X_n)$ 为连续型随机向量即具有联合分布密度 $p(x_1,\cdots,x_n)$, 则

$$E[g(X_1,\cdots,X_n)]=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1,\cdots,x_n)p(x_1,\cdots,x_n)dx_1\cdots dx_n$$

$$E(X_i)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} x_i p(x_1,\cdots,x_n)dx_1\cdots dx_n=\int_{-\infty}^{+\infty} x_i p_{X_i}(x_i)dx_i$$

定义 1.6 假设随机向量 $(X_1, \cdots, X_n)$ 的每个分量 X_i 的数学期望都存在，则称

$$E(X) = (EX_1, EX_2, \cdots, EX_n)$$

为随机向量 X 的数学期望向量，简称为 X 的数学期望。这里

$$E(X_i) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_i dF(x_1, \cdots, x_n) = \int_{-\infty}^{+\infty} x_i dF_i(x_i)$$

这里 $F_i(x_i)$ 为 X_i 的分布函数.

期望的极限性质：单调收敛定理

定理 1.6 (单调收敛定理) 若随机变量序列 $\{X_n\}$ 满足条件

$$0 \le X_1(\omega) \le X_2(\omega) \le \cdots \le X_n(\omega) \uparrow X(\omega), \omega \in \Omega,$$

则

$$\begin{aligned} \int_\Omega \lim_{n\to\infty} X_n(\omega) P(d\omega) &= \lim_{n\to\infty} \int_\Omega X_n(\omega) P(d\omega), \\ \text{i.e.} \quad E(\lim_{n\to\infty} X_n) &= \lim_{n\to\infty} E(X_n). \end{aligned}$$

单调收敛定理的证明 I

- 设 $\{X_{nk}\}_{k\geqslant 1}$ 为定义 $E(X_n)$ 的简单随机变量列，且满足

$$0 \leqslant X_{11} \leqslant X_{12} \leqslant \cdots \leqslant X_{1k} \uparrow X_1$$

$$0 \leqslant X_{21} \leqslant X_{22} \leqslant \cdots \leqslant X_{2k} \uparrow X_2$$

$$\vdots$$

$$0 \leqslant X_{n1} \leqslant X_{n2} \leqslant \cdots \leqslant X_{nk} \uparrow X_n$$

$$\lim_{k\to\infty} E(X_{nk}) = E(X_n)$$

- 令 $\widetilde{X}_k = \max_{1\leqslant i\leqslant k} X_{ik}$ ，则 $\widetilde{X}_k$ 为非降简单随机变量序列，

$$X_{nk} \leqslant \widetilde{X}_k \leqslant X_k, \quad E(X_{nk}) \leq E\left(\widetilde{X}_k\right) \leq E(X_k)$$

- 下证 $\lim_{k\to\infty}\widetilde{X}_k = X$, $\lim_{k\to\infty}E(\widetilde{X}_k) = \lim_{n\to\infty}E(X_n)$
  - 固定 n，令 $k\to+\infty$ 可得

$$X_n \le \lim_{k\to+\infty}\widetilde{X}_k \le X, \qquad E\left(X_n\right) \le \lim_{k\to+\infty} E\left(\widetilde{X}_k\right) \le \lim_{k\to+\infty} E(X_k)$$

  - 再令 $n\to+\infty$ 可得

$$\lim_{k\to\infty}\widetilde{X}_k = X, \qquad \lim_{k\to\infty}E(\widetilde{X}_k) = \lim_{n\to\infty}E(X_n)$$

- 结合 $E(\lim_{n\to\infty}X_n) = E(X) := \lim_{n\to\infty}E(\widetilde{X}_k)$ 可知定理得证.

定理 1.7 (Fatou 引理) 若 $\{X_n\}$ 是一随机变量序列，

- 如果存在可积随机变量 $\underline{Y}$ 使得 $X_n \geq \underline{Y}$, 则有

$$\begin{aligned} \int_\Omega \liminf_{n\to\infty} X_n(\omega) P(d\omega) &\leq \liminf_{n\to\infty} \int_\Omega X_n(\omega) P(d\omega), \\ \text{i.e.} \quad E(\liminf_{n\to\infty} X_n) &\leq \liminf_{n\to\infty} E(X_n). \end{aligned}$$

- 如果存在可积随机变量 $\overline{Y}$ 使得 $X_n \leq \overline{Y}$, 则有

$$\begin{aligned} \int_\Omega \limsup_{n\to\infty} X_n(\omega)) P(d\omega) &\geq \limsup_{n\to\infty} \int_\Omega X_n(\omega)) P(d\omega), \\ \text{i.e.} \quad E(\limsup_{n\to\infty} X_n) &\geq \limsup_{n\to\infty} E(X_n). \end{aligned}$$

Fatou 引理证明

- 如果 $\liminf_{n\to\infty} E(X_n) = +\infty$, 则不等式显然成立. 以下设 $\liminf_{n\to\infty} E(X_n)$ 有限.
- 令 $Y_n = \inf_{k\geq n}(X_k - \underline{Y}), n \geq 1$, 则 $\{Y_n\}$ 为非负随机变量的上升序列，且以 $\lim_n (X_n - \underline{Y})$ 为极限.
- 由单调收敛定理可得:

$$E\left[\liminf_n (X_n - \underline{Y})\right] = \lim_n E\left[\inf_{k\geq n}(X_k - \underline{Y})\right] \leq \liminf_{n\to\infty} E(X_n - \underline{Y}).$$

- 消去有限的 $E(\underline{Y})$ 可证得结论
- 对 $\{-X_n\}$ 用下确界的结论可证得上确界的结论.

定理 1.8 (控制收敛定理) 设 $\{X_n\}$ 是一随机变量序列。如果存在可积随机变量 Y 使得 $|X_n| \le Y$, 且 $\lim_{n\to\infty} X_n(\omega) = X(\omega)$, 则随机变量 $X(\omega)$ 可积，且有

$$\begin{aligned}\int_\Omega \lim_{n\to\infty} X_n(\omega)P(d\omega) &= \lim_{n\to\infty} \int_\Omega X_n(\omega)P(d\omega),\\ \text{i.e.}\quad E(X) = E(\lim_{n\to\infty} X_n) &= \lim_{n\to\infty} E(X_n).\end{aligned}$$

控制收敛定理的证明

- $|X_n| \leq Y$ 蕴含 $|X| \leq Y$, 故 X 可积.
- 取 $\underline{Y} = -Y$ 与 $\overline{Y} = Y$, 用 Fatou 引理便得到

$$E(X) \leq \varliminf_{n} E\left(X_n\right) \leq \varlimsup_{n_0} E\left(X_n\right) \leq E(X)$$

- 故 $\lim_n E\left(X_n\right)$ 存在且等于 $E(X)$, 定理至此得证.

随机变量方差的定义

定义 1.7 若随机变量 X^2 的期望 $E(X^2)$ 存在，则称

$$D(X) := E\left[(X - EX)^2\right]$$

为随机变量 X 的方差。有时我们也用 $Var(X)$ 来表示 X 的方差。称方差 $D(X)$ 的正平方根 $\sqrt{D(X)}$ 为随机变量 X 的标准差，记为 $\sigma(X)$ 或 σ_X.

定理 1.9 假设随机变量 X 的方差存在，则

$$\begin{aligned} D(X) &= E[(X - EX)^2] = \int_{-\infty}^{+\infty} (x - EX)^2 dF(x) \\ &= \begin{cases} \sum_i (x_i - EX)^2 [F(x_i) - F(x_{i-1})] = \sum_i (x_i - EX)^2 P(X = x_i) \\ \\ \int_{-\infty}^{+\infty} (x - EX)^2 p(x) dx \end{cases} \end{aligned}$$

方差的性质

1. $D(X)=E[(X-EX)^2]\geq 0$, $D(X)=0$ 当且仅当 $P(X=EX)=1$;
2. 常数的方差为 0 即 $D(c)=0$, 其中 c 为常数;
3. $D(X)=E(X^2)-(EX)^2$: 事实上，根据方差的定义及期望的线性性质，我们有

$$\begin{aligned} D(X) &= E[(X-EX)^2]=E[X^2-2X\cdot EX+(EX)^2] \\ &= E(X^2)-2EX\cdot EX+(EX)^2 \\ &= E(X^2)-(EX)^2 \end{aligned}$$

4. 若 $E(X^2)=0$, 则 $E(X)=0, D(X)=0$. 事实上

$$0 \le D(X) = E(X^2) - (EX)^2 = -(EX)^2 \le 0$$

5. 若 a,b 为常数, 则 $D(aX+b)=a^2D(X)$, 事实上,

$$\begin{aligned} D(aX+b) &= E(aX+b-E(aX+b))^2 = E(aX-aEX)^2 \\ &= E[a^2(X-EX)^2] = a^2E[(X-EX)^2] = a^2D(X) \end{aligned}$$

6. $f(c) := E(X-c)^2$ 当且仅当 $c=E(X)$ 时取到最小值:由

$$f(c) = E(X-EX+EX-c)^2 = E(X-EX)^2 + [EX-c]^2$$

可得.

切比雪夫 (Chebyshev) 不等式

定理 1.10 设随机变量 X 的期望与方差均存在，则对任意的 $\epsilon > 0$ 有

$$P(|X - EX| \geq \epsilon) \leq \frac{D(X)}{\epsilon^2}$$

证明: 记 $a = EX$, 则

$$\begin{aligned} P(|X-a| \geq \epsilon) &= \int_{\{x:|x-a|\geq\epsilon\}} dF(x) \leq \int_{\{x:|x-a|\geq\epsilon\}} \frac{(x-a)^2}{\epsilon^2} dF(x) \\ &\leq \frac{1}{\epsilon^2} \int_{-\infty}^{+\infty} (x-a)^2 dF(x) = \frac{D(X)}{\epsilon^2} \end{aligned}$$

定理 1.11 若随机变量 X 的方差存在，则

$$D(X)=0 \Leftrightarrow P(X=EX)=1.$$

特别的，如果 $E(X^2)=0$, 则 $E(X)=0, D(X)=0$, 从而 $P(X=0)=1$.

证明: 充分性显然，下面证必要性。设 $D(X)=0$, 此时 EX 存在。注意到

$$\{|X-EX|>0\} = \cup_{n=1}^{\infty}\{|X-EX| \geq \frac{1}{n}\}$$

故

$$\begin{aligned} P(|X-EX|>0) &= P(\cup_{n=1}^{\infty}\{|X-EX| \geq \frac{1}{n}\}) \\ &\leq \sum_{n=1}^{\infty} P(|X-EX| \geq \frac{1}{n}) \leq \sum_{n=1}^{\infty} \frac{D(X)}{(1/n)^2} = 0 \end{aligned}$$

从而

$$P(X=EX) = 1 - P(|X-EX|>0) = 1-0=1$$

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{n} kP(X=k) = \sum_{k=0}^{n} \frac{k\cdot n!}{k!(n-k)!}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!}p^k(1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n\cdot (n-1)!}{(k-1)!(n-1-(k-1))!}pp^{k-1}(1-p)^{n-1-(k-1)} \\
&= np\sum_{k=1}^{n} C_{n-1}^{k-1}p^{k-1}(1-p)^{n-1-(k-1)} \\
&= np
\end{aligned}
$$

特别的，对于两点分布 $(B(1,p))$ 随机变量 X, 有 $E(X)=p$. 事实上对于两点分布其期望为: $E(X)=0\cdot P(X=0)+1\cdot P(X=1)=p$.

计算二项分布期望的另一方法

前面在讲二项分布时，我们知道:

$$X = X_1 + X_2 + \cdots + X_n$$

其中 $X_k, k = 1, \cdots, n$ 为第 k 次伯努利试验成功的次数，显然服从两点分布。故

$$\begin{aligned} EX_i &= 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = p \\ EX_i^2 &= 0^2 \cdot P(X_i = 0) + 1^2 \cdot P(X_i = 1) = p \\ D(X_i) &= EX_i^2 - (EX_i)^2 = p - p^2 = p(1-p) \end{aligned}$$

再由期望的线性性知

$$E(X) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$$

二项分布的方差 $D(X)$

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 P(X=k) = \sum_{k=1}^{n} k(k-1+1) C_n^k p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k(k-1) C_n^k p^k (1-p)^{n-k} + \sum_{k=1}^{n} k C_n^k p^k (1-p)^{n-k} \\
&= \sum_{k=2}^{n} k(k-1) C_n^k p^k (1-p)^{n-k} + np \\
&= n(n-1)p^2 \sum_{k=2}^{n} C_{n-2}^{k-2} p^{k-2} (1-p)^{(n-2)-(k-2)} + np \\
&= n(n-1)p^2 + np
\end{aligned}
$$

故

$$
D(X) = E(X^2) - (EX)^2 = n(n-1)p^2 + np - (np)^2 = np(1-p)
$$

特别的，两点分布的方差为 $p(1-p)$.

泊松分布的期望与方差: $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} kP(X=k) = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} \lambda\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda} = \lambda \\
E(X^2) &= \sum_{k=0}^{\infty} k^2P(X=k) = \sum_{k=1}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty}[(k-1)+1]\frac{\lambda^k}{(k-1)!}e^{-\lambda} \\
&= \lambda^2e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = \lambda^2+\lambda \\
D(X) &= E(X^2)-(EX)^2 = \lambda^2+\lambda-\lambda^2 = \lambda
\end{aligned}
$$

几何分布的期望与方差:$P(X=k)=(1-p)^{k-1}p$

$$
\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k(1-p)^{k-1}p = p\sum_{k=1}^{\infty}\sum_{n=1}^{k}(1-p)^{k-1} = p\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}(1-p)^{k-1} \\
&= p\sum_{n=1}^{\infty}(1-p)^{n-1}\frac{1}{1-(1-p)} = \sum_{n=1}^{\infty}(1-p)^{n-1} = \frac{1}{p} \\
E(X^2) &= \sum_{k=1}^{\infty} k^2(1-p)^{k-1}p = p\left[\sum_{k=1}^{\infty} k(k+1)(1-p)^{k-1} - \sum_{k=1}^{\infty} k(1-p)^{k-1}\right] \\
&= 2p\sum_{k=1}^{\infty}\sum_{n=1}^{k} n(1-p)^{k-1} - p\sum_{k=1}^{\infty} k(1-p)^{k-1} \\
&= 2p\sum_{n=1}^{\infty} n\sum_{k=n}^{\infty}(1-p)^{k-1} - \frac{1}{p} = 2p\sum_{n=1}^{\infty} n(1-p)^{n-1}/p - \frac{1}{p} \\
&= \frac{2}{p^2} - \frac{1}{p} \\
D(X) &= E(X^2) - (EX)^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}
\end{aligned}
$$

负二项分布的期望与方差: $P(X=k)=C_{k-1}^{r-1}p^r(1-p)^{k-r}$

$$
\begin{aligned}
E(X) &= \sum_{k=r}^{\infty} kC_{k-1}^{r-1}p^r(1-p)^{k-r} = \sum_{k=r}^{\infty} \frac{r}{p}\frac{k\cdot(k-1)!}{r\cdot(r-1)!(k-r)!}p^{r+1}(1-p)^{k-r} \\
&= \sum_{k=r}^{\infty} \frac{r}{p} C_{k+1-1}^{r+1-1}p^{r+1}(1-p)^{k-r} = \frac{r}{p}\sum_{l=r+1}^{\infty} C_{l-1}^{r+1-1}p^{r+1}(1-p)^{l-(r+1)} = \frac{r}{p} \\
E(X^2) &= \sum_{k=r}^{\infty} k^2 C_{k-1}^{r-1}p^r(1-p)^{k-r} = \sum_{k=r}^{\infty} k(k+1-1)C_{k-1}^{r-1}p^r(1-p)^{k-r} \\
&= \sum_{k=r}^{\infty} k(k+1)C_{k-1}^{r-1}p^r(1-p)^{k-r} - \sum_{k=r}^{\infty} kC_{k-1}^{r-1}p^r(1-p)^{k-r} \\
&= \sum_{k=r}^{\infty} \frac{r(r+1)}{p^2} C_{k+2-1}^{r+2-1}p^{r+2}(1-p)^{(k+2)-(r+2)} - \frac{r}{p} \\
&= \frac{r(r+1)}{p^2} - \frac{r}{p} \\
D(X) &= E(X^2)-(EX)^2 = \frac{r(r+1)}{p^2} - \frac{r}{p} - \left(\frac{r}{p}\right)^2 = \frac{r(1-p)}{p^2}
\end{aligned}
$$

均匀分布的期望与方差:$X \sim U(a,b)$

- 若 $X \sim U(a,b)$, 即 $p(x) = \begin{cases} \dfrac{1}{b-a}, & x \in (a,b) \\ 0, & 其他 \end{cases}$, 故

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} xp(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{a+b}{2} \\
E(X^2) &= \int_{-\infty}^{+\infty} x^2p(x)dx = \int_a^b x^2\frac{1}{b-a}dx = \frac{a^2+ab+b^2}{3} \\
D(X) &= E(X^2)-(EX)^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}
\end{aligned}
$$

- 若 $X \sim N(0,1)$ 即 $p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ 则

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x p(x) dx = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 \\
D(X) &= E[(X - EX)^2] = E(X^2) = \int_{-\infty}^{+\infty} x^2 p(x) dx \\
&= \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x d(-e^{-\frac{x^2}{2}}) \\
&= \frac{1}{\sqrt{2\pi}} \left(-x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \right) \\
&= 1
\end{aligned}
$$

- 若 $X \sim N(\mu, \sigma^2)$, 则易知 $\dfrac{X-\mu}{\sigma} \sim N(0,1)$, 故

$$
\begin{aligned}
E(X) &= E(\sigma \cdot \frac{X-\mu}{\sigma} + \mu) = \sigma E(\frac{X-\mu}{\sigma}) + \mu = \mu \\
D(X) &= D(\sigma \cdot \frac{X-\mu}{\sigma} + \mu) = \sigma^2 D(\frac{X-\mu}{\sigma}) = \sigma^2
\end{aligned}
$$

Gamma 分布的期望

- 若 $X \sim \Gamma(\alpha, \lambda)$ 即 $p(x) = \begin{cases} \dfrac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$，故

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} x p(x) dx = \int_{0}^{+\infty} x \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\
&= \int_{0}^{+\infty} \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} x^{\alpha+1-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)} = \frac{\alpha}{\lambda}; \\
E(X^2) &= \int_{-\infty}^{+\infty} x^2 p(x) dx = \int_{0}^{+\infty} x^2 \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\
&= \int_{0}^{+\infty} \frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)} \frac{\lambda^{\alpha+2}}{\Gamma(\alpha+2)} x^{\alpha+2-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha+2)}{\lambda^2\Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{\lambda^2} \\
D(X) &= E(X^2) - (EX)^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}
\end{aligned}
$$

指数分布与 χ^2 分布的期望与方差

- 若 $X \sim \Gamma(1, \lambda)$ 分布即 X 服从指数分布，此时 $\alpha = 1$

$$
\begin{aligned}
E(X) &= \frac{\alpha}{\lambda} = \frac{1}{\lambda} \\
E(X^2) &= \frac{\alpha(\alpha+1)}{\lambda^2} = \frac{2}{\lambda^2} \\
D(X) &= \frac{\alpha}{\lambda^2} = \frac{1}{\lambda^2}
\end{aligned}
$$

- 若 $X \sim \Gamma(\frac{n}{2}, \frac{1}{2})$ 即 X 服从 $\chi^2(n)$ 分布，此时 $\alpha = n/2, \lambda = 1/2$

$$
\begin{aligned}
E(X) &= \frac{\alpha}{\lambda} = n \\
E(X^2) &= \frac{\alpha(\alpha+1)}{\lambda^2} = n(n+2) \\
D(X) &= \frac{\alpha}{\lambda^2} = 2n
\end{aligned}
$$

第十八讲：协方差与相关系数

Email: xzhangseu@seu.edu.cn

东南大学　　数学学院

定义 1.8 对于随机向量 $X=(X_1,\cdots,X_n)$, 类似于随机向量期望的定义，我们定义随机向量 X 的方差为 $D(X):=(D(X_1),D(X_2),\cdots,D(X_n))$.

定义 1.9 设随机向量 $(X_1,\cdots,X_n)$ 的每个分量 X_i 的方差均存在，称

$$cov(X_i,X_j):=E[(X_i-EX_i)(X_j-EX_j)],i,j=1,2,\cdots,n$$

为 X_i 与 X_j 的协方差. 而将协方差构成的 $n\times n$ 方阵

$$B=(b_{ij}),\quad b_{ij}=cov(X_i,X_j)$$

称为随机向量的协方差阵。一般来说，

- 若 $cov(X_i,X_j)>0$ 时，称 X_i 与 X_j 正相关;
- 若 $cov(X_i,X_j)<0$ 时，称 X_i 与 X_j 负相关;
- 若 $cov(X_i,X_j)=0$ 时，称 X_i 与 X_j 不相关或零相关.

协方差的性质:$cov(X_i,X_j)=E[(X_i-EX_i)(X_j-EX_j)]$

- $cov(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)$: 事实上

$$
\begin{aligned}
cov(X_i,X_j) &= E[(X_i-EX_i)(X_j-EX_j)] \\
&= E\Big[X_iX_j-X_iE(X_j)-X_jE(X_i)+E(X_i)E(X_j)\Big] \\
&= E(X_iX_j)-E(X_i)E(X_j)
\end{aligned}
$$

- 若 a 为常数，则 $cov(X_i,a)=0$;
- 对任意常数 a,b, 有 $cov(aX_i,bX_j)=ab\cdot cov(X_i,X_j)$;
- $cov(X_i+X_j,X_k)=cov(X_i,X_k)+cov(X_j,X_k)$.

- $cov(X_i, X_i) = E(X_i - EX_i)^2 = D(X_i)$;
- $D(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^n D(X_i) + 2\sum_{1\le i<j\le n} cov(X_i, X_j)$: 事实上

$$
\begin{aligned}
D(\sum_{i=1}^n X_i) &= E(\sum_{i=1}^n X_i - E\sum_{i=1}^n X_i)^2 = E(\sum_{i=1}^n (X_i - EX_i))^2 \\
&= E\Big[\sum_{i=1}^n (X_i - EX_i)^2 + 2\sum_{1\le i<j\le n} (X_i - EX_i)(X_j - EX_j)\Big] \\
&= \sum_{i=1}^n D(X_i) + 2\sum_{1\le i<j\le n} cov(X_i, X_j)
\end{aligned}
$$

- 特别的,

$$D(X_i + X_j) = D(X_i) + D(X_j) + 2cov(X_i, X_j)$$

- 对称性: $cov(X_i, X_j) = cov(X_j, X_i)$, 由定义可直接推得，从而协方差阵 B 是对称矩阵;
- 协方差矩阵 B 是非负定矩阵，事实上，对任意向量 $y = (y_1, y_2, \cdots, y_n)$,

$$
\begin{aligned}
yBy' &= \sum_{i,j} y_i y_j b_{ij} = \sum_{i,j} y_i y_j E[(X_i - EX_i)(X_j - EX_j)] \\
&= E \sum_{i,j} y_i (X_i - EX_i) \cdot y_j (X_j - EX_j) \\
&= E[\sum_i y_i (X_i - EX_i)]^2 \geq 0
\end{aligned}
$$

例 1.1 设二维随机变量 (X,Y) 的联合密度为

$$p(x,y)=\begin{cases}3x, & 0<y<x<1,\\ 0, & \text{其他}\end{cases}$$

试求 $cov(X,Y)$.

解: 由 $cov(X,Y)=E(XY)-EXEY$ 知，我们只需计算 $E(XY),EX,EY$ 的值.

$$\begin{aligned}
EX &= \int_{-\infty}^{+\infty} xp_X(x)dx=\int_{-\infty}^{+\infty} x\int_{-\infty}^{+\infty} p(x,y)dydx\\
&= \int_0^1 x\int_0^x 3xdydx=\int_0^1 3x^3dx=3/4\\
EY &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yp(x,y)dydx=\int_0^1\int_0^x 3xydydx=3/8\\
E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp(x,y)dydx=\int_0^1\int_0^x xy\cdot 3xdydx=3/10\\
cov(X,Y) &= \frac{3}{10}-\frac{3}{4}\cdot\frac{3}{8}=\frac{3}{160}>0
\end{aligned}$$

例 1.2 设二维随机变量 (X,Y) 的联合密度为

$$p(x,y)=\begin{cases}\dfrac{x+y}{3}, & 0<x<1, 0<y<2,\\ 0, & \text{其他}\end{cases}$$

试求 $D(2X-3Y+8)$.

解: 注意到

$$\begin{aligned} D(2X-3Y+8) &= D(2X-3Y)=D(2X)+D(-3Y)+2cov(2X,-3Y)\\ &= 2^2D(X)+(-3)^2D(Y)+2\cdot 2\cdot(-3)cov(X,Y)\\ &= 4D(X)+9D(Y)-12cov(X,Y)\end{aligned}$$

所以我们需要计算:$E(X), E(X^2), E(Y), E(Y^2), E(XY)$.

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xp(x,y)dydx=\int_0^1\int_0^2 x\frac{x+y}{3}dydx=5/9\\ E(X^2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2p(x,y)dydx=\int_0^1\int_0^2 x^2\frac{x+y}{3}dydx=7/18\end{aligned}$$

类似的可计算

$$
\begin{aligned}
E(Y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yp(x,y)dydx = \int_0^1\int_0^2 y\frac{x+y}{3}dydx = 11/9 \\
E(Y^2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2p(x,y)dydx = \int_0^1\int_0^2 y^2\frac{x+y}{3}dydx = 16/9 \\
E(XY) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyp(x,y)dydx = \int_0^1\int_0^2 xy\frac{x+y}{3}dydx = 2/3
\end{aligned}
$$

从而

$$
\begin{aligned}
D(X) &= 7/18-(5/9)^2 = 13/162, \\
D(Y) &= 16/9-(11/9)^2 = 23/81, \\
cov(X,Y) &= 2/3-5/9\cdot 11/9 = -1/81 \\
D(2X-3Y+8) &= 4\cdot 13/162+9\cdot 23/81-12\cdot(-1/81) = 245/81
\end{aligned}
$$

二维正态分布的协方差矩阵

若 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$ 即其联合分布密度为

$$\begin{aligned} p(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ &\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\} \end{aligned}$$

则由上一章的知识知:$X \sim N(\mu_1,\sigma_1^2), Y \sim N(\mu_2,\sigma_2^2)$, 故

$$\begin{aligned} &E(X) = \mu_1, \quad E(Y) = \mu_2 \\ &b_{11} = cov(X,X) = D(X) = \sigma_1^2, \\ &b_{22} = cov(Y,Y) = D(Y) = \sigma_2^2 \end{aligned}$$

$$b_{12} = cov(X,Y) = E[(X-EX)(Y-EY)] = \iint (x-\mu_1)(y-\mu_2)p(x,y)dxdy$$

$$\overset{u=(x-\mu_1)/\sigma_1}{\underset{v=(y-\mu_2)/\sigma_2}{=\!=\!=\!=\!=\!=}} \iint \frac{\sigma_1\sigma_2 uv}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)\}dudv$$

$$= \iint \frac{\sigma_1\sigma_2 uv}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}\left[(u-\rho v)^2+(1-\rho^2)v^2\right]\}dudv$$

$$\overset{s=(u-\rho v)/\sqrt{1-\rho^2}}{\underset{t=v}{=\!=\!=\!=\!=\!=\!=\!=}} \frac{\sigma_1\sigma_2}{2\pi} \iint (\sqrt{1-\rho^2}st+\rho t^2)\exp\{-\frac{s^2+t^2}{2}\}dsdt$$

$$= \frac{\sigma_1\sigma_2}{2\pi}\left(\sqrt{1-\rho^2}\int_{-\infty}^{+\infty} se^{-\frac{s^2}{2}}ds\int_{-\infty}^{+\infty} te^{-\frac{t^2}{2}}dt + \rho\int_{-\infty}^{+\infty} e^{-\frac{s^2}{2}}ds\int_{-\infty}^{+\infty} t^2e^{-\frac{t^2}{2}}dt\right)$$

$$= \frac{\sigma_1\sigma_2}{2\pi}(\sqrt{1-\rho^2}\cdot 0+\rho\sqrt{2\pi}\cdot\sqrt{2\pi}) = \rho\sigma_1\sigma_2$$

故二维正态分布的协方差矩阵为

$$B = \begin{pmatrix} \sigma_1^2, & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2, & \sigma_2^2 \end{pmatrix}$$

定义 1.10 设 X, Y 为方差存在的两个随机变量，则称

$$r := \frac{cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为 X 与 Y 的相关系数.

相关系数的性质: $r := \dfrac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$

- $r = 0 \Leftrightarrow cov(X,Y) = 0$ 即 X, Y 不相关;
- 若记 $X^* := \dfrac{X-EX}{\sqrt{D(X)}}$, $\quad Y^* := \dfrac{Y-EY}{\sqrt{D(Y)}}$, 则 $EX^* = EY^* = 0$, 从而

$$
\begin{aligned}
r &:= \frac{cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{E[(X-EX)(Y-EY)]}{\sqrt{D(X)}\sqrt{D(Y)}} \\
&= E(X^*Y^*) = E[(X^*-EX^*)(Y^*-EY^*)] \\
&= cov(X^*, Y^*)
\end{aligned}
$$

- $|r| \le 1$, 且
  - $r = 1$ 当且仅当 $P(X^* = Y^*) = 1$;
  - $r = -1$ 当且仅当 $P(X^* = -Y^*) = 1$.

定理 1.12 对任意的随机变量 X 与 Y 都有

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

等式成立当且仅当存在常数 t_0 使得

$$P(Y = t_0 X) = 1$$

证明: 对任意的实数 t, 定义

$$u(t) := E(tX - Y)^2 = t^2E(X^2) - 2tE(XY) + E(Y^2).$$

显然，对一切的 $t \in R$, $u(t) \ge 0$, 故方程 $u(t) = 0$ 或者没有实根或者有重根，从而

$$\Delta = [2E(XY)]^2 - 4E(X^2)E(Y^2) \le 0$$

即

$$[E(XY)]^2 \le E(X^2)E(Y^2)$$

上述不等式等号成立，当且仅当 $\Delta = 0$ 即 $u(t) = 0$ 存在一个重根 t_0, 这时

$$u(t_0) = E(t_0X - Y)^2 = 0$$

从而由可知

$$P(t_0X - Y = 0) = 1 \Leftrightarrow P(Y = t_0X) = 1.$$

相关系数绝对值小于等于 1 的证明

- 首先由 $X^* = \dfrac{X-EX}{\sqrt{D(X)}}, Y^* = \dfrac{Y-EY}{\sqrt{D(Y)}}$ 知:

$$\begin{aligned} EX^* &= EY^* = 0 \\ D(X^*) &= D\left(\frac{X-EX}{\sqrt{D(X)}}\right) = \frac{1}{D(X)}D(X-EX) = \frac{1}{D(X)}D(X) = 1 \\ D(Y^*) &= D\left(\frac{Y-EY}{\sqrt{D(Y)}}\right) = \frac{1}{D(Y)}D(Y-EY) = \frac{1}{D(Y)}D(Y) = 1 \end{aligned}$$

- 其次，由 $r = cov(X^*, Y^*)$ 知

$$\begin{aligned} |r| &= |cov(X^*, Y^*)| = |E(X^*Y^*) - E(X^*)E(Y^*)| \leq \sqrt{E[(X^*)^2]E[(Y^*)^2]} \\ &= \sqrt{D(X^*)D(Y^*)} = 1; \end{aligned}$$

- $|r| = 1$ 当且仅当存在 t_0 使得 $P(Y^* = t_0X^*) = 1$, 而此时有

$$r = E(X^*Y^*) = t_0E[(X^*)^2] = t_0$$

定理 1.13 对于随机变量 X, Y, 下面四个事实是等价的:

1. X, Y 不相关;
2. $r = 0$ 即相关系数为 0;
3. $E(XY) = E(X)E(Y)$;
4. $D(X + Y) = D(X) + D(Y)$.

证明: (1) 与 (2) 等价是显然的，根据定义即可得. 由于

$$cov(X, Y) = E(XY) - E(X)E(Y)$$

故 $cov(X, Y) = 0$ 当且仅当 $E(XY) = E(X)E(Y)$, 即 (1) 和 (3) 等价. 另外，又由于

$$D(X + Y) = D(X) + D(Y) + 2cov(X, Y)$$

故 $cov(X, Y) = 0$ 当且仅当 $D(X + Y) = D(X) + D(Y)$, 即 (1) 与 (4) 等价.

独立与不相关的联系：独立必定不相关

定理 1.14 若 X, Y 独立，则 X 与 Y 不相关.

证明: 我们仅对连续型随机变量给出证明。因为 X, Y 独立，故其联合分布密度 $p(x,y) = p_X(x)p_Y(y)$, 从而

$$\begin{aligned} E(XY) &= \iint xyp(x,y)dxdy = \iint xyp_X(x)p_Y(y)dxdy \\ &= \int xp_X(x)dx \int yp_Y(y)dy = E(X)E(Y) \end{aligned}$$

从而 X, Y 不相关.

定理 1.15 若 X, Y 独立，则

$$E(XY) = E(X)E(Y), \quad D(X+Y) = D(X) + D(Y)$$

更一般的，我们有若 $X_1, \cdots, X_n$ 为相互独立的随机变量，则.

$$\begin{aligned} E(X_1 \cdots X_n) &= E(X_1)E(X_2)\cdots E(X_n); \\ D(\sum_{i=1}^n X_i) &= \sum_{i=1}^n D(X_i) \end{aligned}$$

例 1.3 设随机变量 $X \sim N(0,\sigma^2)$, 并令 $Y = X^2$, 显然 X, Y 不独立，但是此时 X, Y 不相关。事实上，此时

$$cov(X,Y) = E(X \cdot X^2) - E(X)E(X^2) = 0.$$

例 1.4 设 θ 为 $[0,2\pi]$ 上的均匀分布，a 为一固定常数，令

$$X=\cos\theta,\quad Y=\cos(\theta+a).$$

则我们有

$$\begin{aligned}
E(X) &= \int_0^{2\pi}\cos t\cdot\frac{1}{2\pi}dt=0,\quad E(Y)=\int_0^{2\pi}\cos(t+a)\cdot\frac{1}{2\pi}dt=0\\
E(X^2) &= \int_0^{2\pi}\cos^2 t\cdot\frac{1}{2\pi}dt=\frac{1}{2},\quad E(Y^2)=\int_0^{2\pi}\cos^2(t+a)\cdot\frac{1}{2\pi}dt=\frac{1}{2}\\
E(XY) &= \int_0^{2\pi}\cos t\cos(t+a)\cdot\frac{1}{2\pi}dt=\frac{1}{2}\cos a
\end{aligned}$$

因此 $r=\cos a$. 故

- $a=0$ 时，$r=1, X=Y$;
- $a=\pi$ 时，$r=-1, X=-Y$;
- $a=\frac{\pi}{2}$ 或 $\frac{3\pi}{2}$ 时，$r=0$, X,Y 不相关，但此时 $X^2+Y^2=1$, 因此不独立.

定理 1.16 对于二维正态分布 (X, Y),X 与 Y 不相关与 X, Y 独立等价.

证明: 因为独立必然不相关，故我们仅需证在不相关条件下，X, Y 独立即可. 对于二维正态分布，我们有

$$cov(X, Y) = \rho\sigma_1\sigma_2$$

从而 X, Y 不相关，当且仅当 $\rho = 0$, 此时

$$\begin{aligned}
p(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\
&\quad \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\} \cdot \frac{1}{\sqrt{2\pi}\sigma_2}\exp\{-\frac{(y-\mu_2)^2}{2\sigma_2^2}\} \\
&= p_X(x)P_Y(y)
\end{aligned}$$

故 X, Y 独立，定理得证.

第十九讲：条件数学期望与母函数

Email: xzhangseu@seu.edu.cn

东南大学　　数学学院

条件数学期望

- 在第二章中对于任何有正概率的事件 B, 我们引入了条件概率 $P(\cdot|B)$ 及条件分布 $F(x|B) := P(X \le x|B)$;
- 类似于概率与分布函数，对于上述的条件概率及条件分布，我们也可以定义相应的条件数学期望:

定义 1.11 如果下述积分绝对收敛，则称

$$E(X|B) := \int_{-\infty}^{+\infty} x dF(x|B)$$

为已知事件 B 发生后 X 的条件数学期望.

定理 1.17 若 X, Y 均为离散型随机变量，并且条件数学期望定义中的 B 选为 $B := \{Y = y_j\}$ 的形式，则

$$E(X|Y = y_j) = \int_{-\infty}^{+\infty} x dF(x|Y = y_j) = \sum_i x_i P(X = x_i|Y = y_j)$$

定理 1.18 若 X, Y 均为连续型随机变量，并且条件数学期望定义中的 B 选为 $B := \{Y = y\}$ 的形式，则

$$E(X|Y = y) = \int_{-\infty}^{+\infty} x dF(x|Y = y) = \int_{-\infty}^{+\infty} x p(x|y) dx$$

定理 1.19 设 $g(x)$ 为 Borel 函数，则 $g(X)$ 关于 $Y=y$ 的条件期望为

$$E(g(X)|Y=y)=\int_{-\infty}^{+\infty}g(x)dF(x|y)$$

定义 1.12 假设 X 的方差存在，则称

$$\begin{aligned}D(X|Y=y) \quad &:= \quad E[(X-E(X|Y=y))^2|Y=y] \\ &= \quad \int_{-\infty}^{+\infty}(x-E(X|Y=y))^2dF(x|y)\end{aligned}$$

为给定 $Y=y$ 后 X 的条件方差.

例 1.5 若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则

$$E(X|Y=y)=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2).$$

证明: 在第二章中，我们知道给定 $Y=y$ 下，X 的条件分布服从

$$N(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\sigma_1^2(1-\rho^2))$$

故

$$E(X|Y=y)=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$$

一般来说，给定 $Y=y$ 后 X 的条件数学期望是 Y 的可能值 y 的函数，记之为

$$\varphi(y):=E(X|Y=y)$$

如果再将 y 用 Y 代回，就得到一个随机变量 $\varphi(Y)$, 相应的，我们记

$$E(X|Y):=\varphi(Y)$$

并称随机变量 $E(X|Y)$ 为 X 关于 Y 的条件数学期望.

条件数学期望的性质：重期望

定理 1.20 设 (X,Y) 为二维随机向量，则对于 Borel 函数 $g(x)$ 有

$$E[E(g(X)|Y)] = E(g(X))$$

证明: 我们仅对二维连续型随机变量给出证明。令
$\varphi(y) := E(g(X)|Y=y)$, 则 $E(g(X)|Y) = \varphi(Y)$, 从而

$$\begin{aligned}
\varphi(y) &= E(g(X)|Y=y) = \int_{-\infty}^{+\infty} g(x)dF(x|y) \\
&= \int_{-\infty}^{+\infty} g(x)p(x|Y=y)dx = \int_{-\infty}^{+\infty} g(x)\frac{p(x,y)}{p_Y(y)}dx \\
E[E(g(X)|Y)] &= E(\varphi(Y)) = \int_{-\infty}^{+\infty} \varphi(y)p_Y(y)dy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)\frac{p(x,y)}{p_Y(y)}p_Y(y)dxdy \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)p(x,y)dxdy = E(g(X))
\end{aligned}$$

重期望的不同表达形式

- 上述定理的重期望公式可写成如下形式

$$
\begin{aligned}
E(g(X)) &= \int_{-\infty}^{+\infty} E[g(X)|Y=y]dF_Y(y) \\
&= \begin{cases} \sum_j E(g(X)|Y=y_j)P(Y=y_j) \\ \int_{-\infty}^{+\infty} E(g(X)|Y=y)p_Y(y)dy \end{cases}
\end{aligned}
$$

- 特别的，若 $g(X)=X$, 则有

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} E[X|Y=y]dF_Y(y) \\
&= \begin{cases} \sum_j E(X|Y=y_j)P(Y=y_j) \\ \int_{-\infty}^{+\infty} E(X|Y=y)p_Y(y)dy \end{cases}
\end{aligned}
$$

例 1.6 一个窃贼被关在有 3 个门的地牢中。其中 1 号门通向自由，出 1 号门后走 3 个小时便回到地面；2 号门通向一个地道，在此地道走 5 个小时后返回地牢；3 号门通向一个更长的地道，沿这个地道走 7 个小时后回到地牢。如果窃贼每次选择 3 个门的可能性总相等，求他为获自由而奔走的平均时间.

解: 设窃贼需要走 X 小时到达地面，则 X 的所有可能取值为

$$3, 5+3, 7+3, 5+5+3, 5+7+3, 7+7+3, \cdots,$$

则显然要写出 X 的分布列是困难的，所以无法直接求 $E(X)$. 但是如果我们引入 Y 表示窃贼每次对 3 个门的选择，则

$$P(Y=1)=P(Y=2)=P(Y=3)=1/3$$
$$E(X|Y=1)=3, E(X|Y=2)=5+E(X), E(X|Y=3)=7+E(X)$$

从而由重期望公式可得

$$E(X)=\sum_{i=1}^{3}E(X|Y=i)P(Y=i)=5+\frac{2}{3}E(X)\Rightarrow E(X)=15$$

定理 1.21 设 $X_1,X_2,\cdots,$ 为一列独立同分布的随机变量，随机变量 N 只取正整数值，且 N 与 $\{X_n\}$ 独立，则

$$E\left(\sum_{i=1}^{N}X_i\right)=E(X_1)E(N)$$

证明: 由重期望公式可知

$$\begin{aligned}
E\left(\sum_{i=1}^{N}X_i\right) &= E\left[E\left(\sum_{i=1}^{N}X_i|N\right)\right]=\sum_{n=1}^{+\infty}E\left(\sum_{i=1}^{N}X_i|N=n\right)P(N=n)\\
&= \sum_{n=1}^{+\infty}E\left(\sum_{i=1}^{n}X_i|N=n\right)P(N=n)=\sum_{n=1}^{+\infty}E\left(\sum_{i=1}^{n}X_i\right)P(N=n)\\
&= \sum_{n=1}^{+\infty}nE(X_1)P(N=n)=E(X_1)E(N)
\end{aligned}$$

定义 1.13 如果 $E|X|^k < +\infty$, 则称 $m_k := E(X^k)$ 为随机变量 X(及其分布) 的 k 阶原点矩。而称 $c_k := E(X-EX)^k$ 为随机变量 (及其分布) 的 k 阶中心矩.

定理 1.22 当 $E|X|^k < +\infty$ 时有,

$$c_k = \sum_{i=0}^{k} C_k^i (-m_1)^{k-i} m_i, \quad m_k = \sum_{i=0}^{k} C_k^i c_{k-i} m_1^i.$$

证明: 注意到

$$\begin{aligned}
(X-EX)^k &= \sum_{i=0}^{k} C_k^i X^i (-EX)^{k-i}, \\
X^k = (EX + X - EX)^k &= \sum_{i=0}^{k} C_k^i (EX)^i (X-EX)^{k-i}
\end{aligned}$$

对上面两式取期望即可得证.

定义 1.14 对任何实数列 $\{p_n\}$, 如果幂级数

$$G(s) = \sum_{n=0}^{\infty} p_n s^n \tag{1}$$

的收敛半径 $s_0 > 0$, 则称 $G(s)$ 为数列 $\{p_n\}$ 的母函数。特别当 $\{p_n\}$ 为某非负整值随机变量 X 的概率分布时，(1) 式至少在区间 $[-1, 1]$ 上绝对收敛且一致收敛，此时有

$$G(s) = E(s^X),$$

称此 $G(s)$ 为随机变量 X 或其概率分布 $\{p_n\}$ 的母函数.

母函数与分布之间的反演公式

- 已知 $G(s)$, 如何确定 $\{p_n\}$?
- 注意到

$$G(0) = p_0, \quad G^{(n)}(s) = n!p_n + \sum_{k=n+1}^{\infty} k(k-1)\cdots(k-n+1)p_k s^{k-n}$$

- 故

$$p_n = \frac{1}{n!}G^{(n)}(0)$$

- 非负整值概率分布 $\{p_n\}$ 与其母函数 $G(s)$ 是一一对应的，母函数可以作为描述这种分布的一种工具.

定理 1.23 设非负随机变量 X 的母函数为 $G(s)$，如果 $E(X)$ 与 $E(X^2)$ 有限，则

$$G'(1) = E(X), \quad G''(1) = E(X^2) - E(X)$$

常见离散型分布的母函数

- Poisson 分布的母函数

$$G(s) = E(s^X) = \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda(s-1)}$$

- 几何分布的母函数

$$G(s) = E(s^X) = \sum_{k=1}^{\infty} s^k q^{k-1} p = \frac{ps}{1-qs}$$

独立和的母函数

定理 1.24 如果 X, Y 为相互独立的随机变量，它们分别有概率分布 $\{a_n\}, \{b_n\}$ 及对应的母函数 $A(s), B(s)$, 则它们的和 $X+Y$ 的母函数为

$$C(s) = A(s)B(s).$$

更进一步，如果 $X_1, \cdots, X_n$ 相互独立，其对应的母函数分别为 $A_1(s)$, $\cdots, A_n(s)$, 则 $X_1 + \cdots + X_n$ 的母函数为

$$C(s) = \Pi_{i=1}^{n} A_i(s)$$

证明： 注意到由 X, Y 相互独立可知 s^X, s^Y 也相互独立，从而

$$C(s) = E(s^{X+Y}) = E(s^X s^Y) = E(s^X)E(s^Y) = A(s)B(s)$$

例 1.7 设 X 服从二项分布 $B(n, p)$, 则 X 的母函数为 $G(s) = (q + ps)^n$.

定理 1.25 设 $\{X_k\}$ 为相互独立同分布的非负整值随机变量序列，其共同的母函数为 $G(s)$. 如果 N 为另一非负整值随机变量，其母函数为 $F(s)$. 则当 N 与每一个 X_k 均独立时，$X=\sum_{k=1}^{N} X_k$ 的母函数为

$$H(s)=F(G(s)).$$

证明: 由重期望公式可得

$$\begin{aligned}
H(s) &= E(s^X)=E(E(s^X|N))=\sum_{n=0}^{\infty}P(N=n)E(s^{\sum_{k=1}^{N}X_k}|N=n)\\
&= \sum_{n=0}^{\infty}P(N=n)E(s^{\sum_{k=1}^{n}X_k}|N=n)=\sum_{n=0}^{\infty}P(N=n)E(s^{\sum_{k=1}^{n}X_k})\\
&= \sum_{n=0}^{\infty}P(N=n)[G(s)]^n=E(G(s)^N)\\
&= F(G(s))
\end{aligned}$$

随机个非负整值随机量之和的期望与方差

- 对 $H(s)$ 求导可得
$$H'(s) = F'(G(s))G'(s)$$
- 令 $s=1$ 并注意到 $G(1)=1$ 可得
$$E(X) = E(N)E(X_1)$$
- 再对 $H'(s)$ 求导可得
$$H''(s) = F''(G(s))[G'(s)]^2 + F'(G(s))G''(s)$$
- 令 $s=1$ 可得
$$\begin{aligned} E(X^2) - E(X) &= [E(N^2) - E(N)][E(X_1)]^2 + E(N)[E(X_1^2) - E(X_1)] \\ &= E(N)D(X_1) + E(N^2)[E(X_1)]^2 - E(N)E(X_1) \end{aligned}$$
- 从而
$$D(X) = E(X^2) - [E(X)]^2 = E(N)D(X_1) + D(N)[E(X_1)]^2$$

第二十讲：特征函数

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

定义 1.15 若 $X(\omega), Y(\omega)$ 为定义在 Ω 上的实值随机变量，则称 $Z(\omega) = X(\omega) + iY(\omega)$ 为复随机变量；称 $\overline{Z} = X(\omega) - iY(\omega)$ 为 $Z(\omega)$ 的复共轭随机变量；称 $|Z| := \sqrt{X^2 + Y^2}$ 为复随机变量 Z 的模.

定义 1.16 若随机变量 X, Y 的期望 $E(X), E(Y)$ 都存在，则复随机变量 Z 的数学期望定义为 $E(Z) := E(X) + iE(Y)$.

- $Z_1 = X_1 + iY_1$ 与 $Z_2 = X_2 + iY_2$ 独立当且仅当 (X_1, Y_1) 与 (X_2, Y_2) 独立;
- $E(e^{iX}) = E(\cos X) + iE(\sin X)$;
- $|e^{iX}| = \sqrt{\cos^2 X + \sin^2 X} = 1$;
- 若 X, Y 独立，则 e^{iX} 与 e^{iY} 也独立.

特征函数的定义

定义 1.17 设 X 是一个随机变量，称

$$\begin{aligned}\varphi(t) &:= E(e^{itX}) = E[\cos(tX)] + iE[\sin(tX)] \\ &= \int_{-\infty}^{+\infty}\cos(tx)dF(x) + i\int_{-\infty}^{+\infty}\sin(tx)dF(x) \\ &= \int_{-\infty}^{+\infty}e^{itx}dF(x), \quad -\infty < t < +\infty\end{aligned}$$

为 X 的特征函数。特别的，

- 如果 X 为离散型随机变量，其分布列为 $p_k = P(X = x_k)$, 则

$$\varphi(t) = \sum_{k=1}^{+\infty} e^{itx_k}p_k, \quad -\infty < t < +\infty;$$

- 如果 X 为连续型随机变量，其分布密度为 $p(x)$, 则

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx}p(x)dx, \quad -\infty < t < +\infty;$$

特征函数的性质 I

1. 由 $\varphi(t)=E[\cos(tX)]+iE[\sin(tX)]$ 知，特征函数总是存在的;
2. $\varphi(0)=1, |\varphi(t)|\leq 1$: $\varphi(0)=E(1)=1$, 而

$$
\begin{aligned}
|\varphi(t)| &= |E[\cos(tX)]+iE[\sin(tX)]| = \left[(E[\cos(tX)])^2+(E[\sin(tX)]\right]^{1/2} \\
&\leq \left[E[\cos^2(tX)]+E[\sin^2(tX)]\right]^{1/2} = 1;
\end{aligned}
$$

3. $\varphi(-t)=\overline{\varphi(t)}$: $\varphi(-t)=E[\cos(tX)]-iE[\sin(tX)]=\overline{\varphi(t)}$;
4. 若 $Y=aX+b$, 其中 a,b 为常数，则 $\varphi_Y(t)=e^{ibt}\varphi_X(at)$, 事实上，

$$
\begin{aligned}
&\varphi_Y(t)=E[e^{it(aX+b)}]=E[\cos(t(aX+b))]+iE[\sin(t(aX+b))] \\
&=E[\cos(taX)\cos(tb)-\sin(taX)\sin(tb)]+iE[\sin(taX)\cos(tb)+\cos(taX)\sin(tb)] \\
&= \cos(tb)\big(E[\cos(taX)]+iE[\sin(taX)]\big)+i\sin(tb)\big(E[\cos(taX)]+iE[\sin(taX)]\big) \\
&=e^{itb}Ee^{itaX}=e^{itb}\varphi_X(at)
\end{aligned}
$$

5. 若 X, Y 独立，则 $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$, 事实上,

$$
\begin{aligned}
&\varphi_{X+Y}(t) = E(e^{it(X+Y)}) = E(e^{itX}e^{itY}) \\
&= E[(\cos(tX) + i\sin(tX))(\cos(tY) + i\sin(tY))] \\
&= E[(\cos(tX) + i\sin(tX))\cos(tY)] + iE[(\cos(tX) + i\sin(tX))\sin(tY)] \\
&= E[\cos(tY)]E[e^{itX}] + iE[\sin(tY)]E[e^{itX}] \\
&= E(e^{itX})E(e^{itY})
\end{aligned}
$$

6. 更一般的，若 $X_1, X_2, \cdots, X_n$ 相互独立，则对于 $Y = X_1 + \cdots + X_n$ 的特征函数有

$$\varphi_Y(t) = \Pi_{k=1}^{n}\varphi_{X_k}(t)$$

7. 如果 $E(X^k)$ 存在，则

$$\varphi^{(k)}(t) = i^k E[X^k e^{itX}], \quad \varphi^{(k)}(0) = i^k E[X^k]$$

事实上，

$$\frac{d^k}{dt^k}\varphi(t) = \frac{d^k}{dt^k}E[e^{itX}] = E[\frac{d^k}{dt^k}e^{itX}] = E[(iX)^k e^{itX}] = i^k E[X^k e^{itX}]$$

8. $\varphi(t)$ 在 $(-\infty,+\infty)$ 上一致连续:

$$
\begin{aligned}
&|\varphi(t+h)-\varphi(t)| = |E(e^{i(t+h)X})-E(e^{itX})| = |E[e^{itX}(e^{ihX}-1)]| \\
&\le E|e^{itX}(e^{ihX}-1)| = E|e^{ihX}-1| = \int_{-\infty}^{+\infty}|e^{ihx}-1|dF_X(x) \\
&= \int_{|x|>A}|e^{ihx}-1|dF_X(x) + \int_{-A}^{A}|e^{ihx}-1|dF_X(x) \\
&\le 2\int_{|x|>A}dF_X(x) + \int_{-A}^{A}|hx|dF_X(x) \\
&\le 2P(|X|>A) + hAP(|X|\le A) \le 2P(|X|>A) + hA
\end{aligned}
$$

选取 A 使得 $P(|X|>A)\le \epsilon/4$, $h\le\delta=\dfrac{\epsilon}{2A}$, 则只要 $|h|\le\delta$ 必有

$$
|\varphi(t+h)-\varphi(t)| \le 2P(|X|>A) + hA \le \epsilon
$$

9. $\varphi(t)$ 是非负定函数，即对任意的正整数 n 及 n 个实数 $t_1, \cdots, t_n$ 及 n 个复数 $z_1, \cdots, z_n$，均有

$$\sum_{k=1}^{n}\sum_{j=1}^{n}\varphi(t_k - t_j)z_k\bar{z}_j \geq 0$$

证明: 由特征函数的定义可得

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{n}\varphi(t_k - t_j)z_k\bar{z}_j &= \sum_{k=1}^{n}\sum_{j=1}^{n} z_k\bar{z}_j E[e^{i(t_k-t_j)X}] = \sum_{k=1}^{n}\sum_{j=1}^{n} E[z_k e^{it_kX}\bar{z}_j\overline{e^{it_jX}}] \\ &= E[\sum_{k=1}^{n} z_k e^{it_kX}\sum_{j=1}^{n}\overline{z_j e^{it_jX}}] = E[|\sum_{k=1}^{n} z_k e^{it_kX}|^2] \geq 0\end{aligned}$$

常用分布的特征函数 I

- 单点分布 $P(X=a)=1$: $\varphi(t)=e^{ita}$;
- 两点分布 $P(X=x)=p^x(1-p)^{1-x}, x=0,1$: $\varphi(t)=pe^{it}+(1-p)$;
- 泊松分布:$\varphi(t)=\sum_{k=0}^{+\infty}e^{ikt}\dfrac{\lambda^k}{k!}e^{-\lambda}=e^{\lambda e^{it}}e^{-\lambda}=e^{\lambda(e^{it}-1)}$;
- 均匀分布 $U(a,b)$:$\varphi(t)=\int_a^b\dfrac{e^{itx}}{b-a}dx=\dfrac{e^{ibt}-e^{iat}}{it(b-a)}$;
- 标准正态分布:

$$
\begin{aligned}
\varphi(t) &= \int_{-\infty}^{+\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\
&= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{+\infty}\cos(tx)e^{-\frac{x^2}{2}}dx + i\int_{-\infty}^{+\infty}\sin(tx)e^{-\frac{x^2}{2}}dx\right) \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\cos(tx)e^{-\frac{x^2}{2}}dx
\end{aligned}
$$

故

$$\begin{aligned}\varphi'(t) &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\sin(tx)e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sin(tx)de^{-\frac{x^2}{2}} \\ &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t\cos(tx)e^{-\frac{x^2}{2}}dx = -t\varphi(t)\end{aligned}$$

从而

$$\frac{d}{dt}[\varphi(t)e^{t^2/2}] = [\varphi'(t) + t\varphi(t)]e^{t^2/2} = 0$$

故 $\varphi(t)e^{t^2/2} = C$, 再利用 $C = \varphi(0) = 1$ 得

$$\varphi(t) = e^{-t^2/2}.$$

常用分布的特征函数 III

- 指数分布:

$$\varphi(t) = \int_0^\infty e^{itx} \lambda e^{-\lambda x} dx = \left(1 - \frac{it}{\lambda}\right)^{-1}$$

- 二项分布: $Y \sim B(n,p)$. 注意到 $Y = X_1 + \cdots + X_n$, 其中 X_i 为独立同分布的随机变量, 且 $X_i \sim B(1,p)$, 故

$$\varphi_Y(t) = \Pi_{k=1}^n (pe^{it} + q) = (pe^{it} + q)^n, \quad q = 1 - p$$

- 一般正态分布: $Y \sim N(\mu, \sigma^2)$, 则由 $X := \dfrac{Y - \mu}{\sigma} \sim N(0,1)$ 及 $Y = \sigma X + \mu$ 可得

$$\varphi_Y(t) = e^{i\mu t} \varphi_X(\sigma t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$$

常用分布的特征函数 IV

- 伽玛分布:$Y \sim \Gamma(n, \lambda)$. 注意到 $Y = X_1 + \cdots + X_n$, 其中 X_i 独立同指数分布即 $X_i \sim \exp(\lambda)$, 故

$$\varphi_Y(t) = \Pi_{k=1}^{n}(1 - \frac{it}{\lambda})^{-1} = (1 - \frac{it}{\lambda})^{-n}$$

  更进一步，当 $Y \sim \Gamma(\alpha, \lambda)$, 我们有

$$\varphi_Y(t) = (1 - \frac{it}{\lambda})^{-\alpha}$$

- $\chi^2(n)$ 分布：因为 $\chi^2(n) = \Gamma(n/2, 1/2)$, 故其特征函数为

$$\varphi(t) = (1 - 2it)^{-n/2}$$

例 1.8 利用特征函数的方法求 $\Gamma(\alpha,\lambda)$ 分布的数学期望与方差.

解: Γ 分布的特征函数及其导数分别为

$$
\begin{aligned}
\varphi(t) &= (1-\frac{it}{\lambda})^{-\alpha} \\
\varphi'(t) &= \frac{i\alpha}{\lambda}(1-\frac{it}{\lambda})^{-\alpha-1}, \quad \varphi'(0)=\frac{i\alpha}{\lambda} \\
\varphi''(t) &= \frac{i^2\alpha(\alpha+1)}{\lambda^2}(1-\frac{it}{\lambda})^{-\alpha-2}, \quad \varphi''(0)=-\frac{\alpha(\alpha+1)}{\lambda^2}
\end{aligned}
$$

故由 $\varphi^{(k)}(0)=i^kE(X^k)$ 知,

$$
\begin{aligned}
E(X) &= \frac{\varphi'(0)}{i}=\frac{\alpha}{\lambda}, \quad E(X^2)=\frac{\varphi''(0)}{i^2}=\frac{\alpha(\alpha+1)}{\lambda^2} \\
D(X) &= E(X^2)-(EX)^2=\frac{\alpha(\alpha+1)}{\lambda^2}-\frac{\alpha^2}{\lambda^2}=\frac{\alpha}{\lambda^2}
\end{aligned}
$$

定理 1.26 (逆转公式) 设 $F(x)$ 与 $\varphi(t)$ 分别为随机变量 X 的分布函数和特征函数，则对 $F(x)$ 的任意两个连续点 $x_1 < x_2$ 有

$$F(x_2) - F(x_1) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi(t) dt$$

定理 1.27 (唯一性定理) 随机变量的分布函数由其特征函数唯一决定.

证明: 在逆转公式中令 $x_2 = x, x_1 = y \to -\infty$ 可得

$$F(x) = \lim_{y\to-\infty} \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ity} - e^{-itx}}{it} \varphi(t) dt$$

而由分布函数本身是右连续可知，分布函数任一点的取值可由连续点上的值唯一决定.

引理 1.1 (狄利克雷积分) 令

$$I(a,T)=\int_0^T \frac{\sin at}{t}dt, \quad \mathrm{sgn}\{a\}:=\begin{cases} -1, & a<0, \\ 0, & a=0, \\ 1, & a>0, \end{cases}$$

则

$$\lim_{T\to+\infty} I(a,T)=\lim_{T\to+\infty}\int_0^T \frac{\sin at}{t}dt=\frac{\pi}{2}\mathrm{sgn}\{a\}$$

证明: 注意到如果作积分换元 $y=at$ 可得

$$\lim_{T\to+\infty} I(a,T)=\mathrm{sgn}\{a\}\lim_{T\to+\infty} I(1,T)$$

只需证明

$$\lim_{T\to+\infty} I(1,T) = \lim_{T\to+\infty}\int_0^T \frac{\sin t}{t}dt = \frac{\pi}{2}.$$

注意到 $\frac{1}{t} = \int_0^{+\infty} e^{-ut}du$, 故

$$\begin{aligned} I(1,T) &= \int_0^T\int_0^{+\infty} e^{-ut}\sin t du dt = \int_0^{+\infty}\left[\int_0^T e^{-ut}\sin t dt\right]du \\ &= \int_0^{+\infty}\left[\frac{1}{1+u^2} - \frac{1}{1+u^2}e^{-uT}(u\sin T + \cos T)\right]du \\ &= \frac{\pi}{2} - \int_0^{+\infty}\frac{u\sin T+\cos T}{1+u^2}e^{-uT}du \to \frac{\pi}{2} \end{aligned}$$

证明: 考虑积分

$$
\begin{aligned}
J(T) &= \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx_1}-e^{-itx_2}}{it}\varphi(t)dt \\
&= \frac{1}{2\pi}\int_{-T}^{T}\left[\int_{-\infty}^{+\infty}\frac{e^{-itx_1}-e^{-itx_2}}{it}e^{itx}dF(x)\right]dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-T}^{T}\frac{e^{it(x-x_1)}-e^{it(x-x_2)}}{it}dt\right]dF(x) \\
&= \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[\int_{0}^{T}\frac{\sin t(x-x_1)}{t}dt-\int_{0}^{T}\frac{\sin t(x-x_2)}{t}dt\right]dF(x) \\
&= \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[I(x-x_1,T)-I(x-x_2,T)\right]dF(x) \\
&\to \frac{1}{\pi}\frac{\pi}{2}\int_{-\infty}^{+\infty}\left[\operatorname{sgn}\{x-x_1\}-\operatorname{sgn}\{x-x_2\}\right]dF(x) \qquad T\to+\infty \\
&= \int_{x_1}^{x_2}dF(x)=F(x_2)-F(x_1)
\end{aligned}
$$

定理 1.28 (逆转公式一般情形) 设 $F(x)$ 与 $\varphi(t)$ 分别为随机变量 X 的分布函数和特征函数，则对任意的 $x_1, x_2 \in R$ 有

$$\frac{F(x_2-0)+F(x_2)}{2}-\frac{F(x_1-0)+F(x_1)}{2}=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx_1}-e^{-itx_2}}{it}\varphi(t)dt$$

定理 1.29 若 X 为连续型随机变量，其密度函数为 $p(x)$, 其特征函数为

$$\varphi(t)=\int_{-\infty}^{+\infty}e^{itx}p(x)dx,$$

如果 $\int_{-\infty}^{+\infty}|\varphi(t)|dt<+\infty$, 则

$$p(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt$$

证明: 由分布密度的定义可知

$$\begin{aligned}
p(x) &= \lim_{h\to 0}\frac{F(x+h)-F(x)}{h}=\lim_{h\to 0}\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{-itx}-e^{-it(x+h)}}{it\cdot h}\varphi(t)dt\\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\lim_{h\to 0}\frac{e^{-ith}-1}{-it\cdot h}e^{-itx}\varphi(t)dt=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}\varphi(t)dt
\end{aligned}$$