# 概率论与数理统计（5）

清华大学

2020 年春季学期

# 二维随机变量函数的数学期望

- 若二维随机变量 $(X, Y)$ 的联合分布列为 $P(X = x_i, Y = y_j)$ 或者联合密度函数为 $p(x, y)$, 则 $Z = g(X, Y)$ 的数学期望为

$$E(Z) = E(g(X, Y)) = \begin{cases} \sum_i \sum_j g(x_i, y_j) P(X = x_i, Y = y_j), \\ \iint g(x, y) p(x, y) \, dx dy. \end{cases}$$

# 二维随机变量函数的数学期望

- 若二维随机变量 $(X, Y)$ 的联合分布列为 $P(X = x_i, Y = y_j)$ 或者联合密度函数为 $p(x, y)$, 则 $Z = g(X, Y)$ 的数学期望为

$$E(Z) = E(g(X, Y)) = \begin{cases} \sum_i \sum_j g(x_i, y_j) P(X = x_i, Y = y_j), \\ \iint g(x, y) p(x, y) \, dx dy. \end{cases}$$

- 或者，我们已经找到 $Z = g(X, Y)$ 的分布列或密度函数，则

$$E(Z) = \begin{cases} \sum_k z_k P(Z = z_k) \\ \int_{-\infty}^{\infty} z p(z) \, dz. \end{cases}$$

# 二维随机变量函数的数学期望

- 若二维随机变量 $(X, Y)$ 的联合分布列为 $P(X = x_i, Y = y_j)$ 或者联合密度函数为 $p(x, y)$, 则 $Z = g(X, Y)$ 的数学期望为

$$E(Z) = E(g(X, Y)) = \begin{cases} \sum_i \sum_j g(x_i, y_j) P(X = x_i, Y = y_j), \\ \iint g(x, y) p(x, y)\, dx dy. \end{cases}$$

- 或者，我们已经找到 $Z = g(X, Y)$ 的分布列或密度函数，则

$$E(Z) = \begin{cases} \sum_k z_k P(Z = z_k) \\ \int_{-\infty}^{\infty} z p(z)\, dz. \end{cases}$$

- 若 $g(X, Y) = X$, 得到的是 $X$ 边际分布的期望:
$E(X) = \int_{-\infty}^{\infty} x p_X(x)\, dx.$

# 二维随机变量函数的数学期望

- 若二维随机变量 $(X, Y)$ 的联合分布列为 $P(X = x_i, Y = y_j)$ 或者联合密度函数为 $p(x, y)$, 则 $Z = g(X, Y)$ 的数学期望为

$$E(Z) = E(g(X, Y)) = \begin{cases} \sum_i \sum_j g(x_i, y_j) P(X = x_i, Y = y_j), \\ \iint g(x, y) p(x, y) \, dxdy. \end{cases}$$

- 或者，我们已经找到 $Z = g(X, Y)$ 的分布列或密度函数，则

$$E(Z) = \begin{cases} \sum_k z_k P(Z = z_k) \\ \int_{-\infty}^{\infty} z p(z) \, dz. \end{cases}$$

- 若 $g(X, Y) = X$, 得到的是 $X$ 边际分布的期望:
$E(X) = \int_{-\infty}^{\infty} x p_X(x) \, dx.$
- 若 $g(X, Y) = (X - E(X))^2$, 得到的是 $X$ 边际分布的方差。

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度:

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度：$X, Y$ 服从 $(0, a)$ 上的均匀分布，且相互独立，

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度：$X$, $Y$ 服从 $(0, a)$ 上的均匀分布，且相互独立，

$$
\begin{aligned}
E(|X-Y|) &= \int_0^a \int_0^a |x-y| \frac{1}{a^2} dxdy \\
&= \frac{1}{a^2} \Big[ \int_0^a \int_0^x (x-y) dydx + \int_0^a \int_x^a (y-x) dydx \Big] \\
&= \frac{1}{a^2} \int_0^a (x^2 - ax + \frac{a^2}{2}) dx = \frac{a}{3}.
\end{aligned}
$$

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度：$X$, $Y$ 服从 $(0,a)$ 上的均匀分布，且相互独立，

$$
\begin{aligned}
E(|X-Y|) &= \int_0^a \int_0^a |x-y| \frac{1}{a^2}\, dxdy \\
&= \frac{1}{a^2}\Big[\int_0^a \int_0^x (x-y)\, dydx + \int_0^a \int_x^a (y-x)\, dydx\Big] \\
&= \frac{1}{a^2}\int_0^a (x^2 - ax + \frac{a^2}{2})\, dx = \frac{a}{3}.
\end{aligned}
$$

- $X_i \sim Exp(\lambda)$, $i=1,2$, 且相互独立，$Y=\max\{X_1, X_2\}$.

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度：$X$, $Y$ 服从 $(0,a)$ 上的均匀分布，且相互独立，

$$
\begin{aligned}
E(|X-Y|) &= \int_0^a \int_0^a |x-y| \frac{1}{a^2}\, dxdy \\
&= \frac{1}{a^2}\Big[\int_0^a \int_0^x (x-y)\, dydx + \int_0^a \int_x^a (y-x)\, dydx\Big] \\
&= \frac{1}{a^2}\int_0^a (x^2 - ax + \frac{a^2}{2})\, dx = \frac{a}{3}.
\end{aligned}
$$

- $X_i \sim Exp(\lambda)$, $i=1,2$, 且相互独立，$Y=\max\{X_1,X_2\}$.
- $p_Y(y) = 2(1-e^{-\lambda y})\lambda e^{-\lambda y}$, $y>0$.

# 例子

- 在长度为 $a$ 的线段上任取两点的平均长度：$X$, $Y$ 服从 $(0, a)$ 上的均匀分布，且相互独立，

$$
\begin{aligned}
E(|X-Y|) &= \int_0^a \int_0^a |x-y| \frac{1}{a^2} dxdy \\
&= \frac{1}{a^2}\Big[\int_0^a \int_0^x (x-y)\,dydx + \int_0^a \int_x^a (y-x)\,dydx\Big] \\
&= \frac{1}{a^2}\int_0^a (x^2 - ax + \frac{a^2}{2})\,dx = \frac{a}{3}.
\end{aligned}
$$

- $X_i \sim Exp(\lambda)$, $i = 1, 2$, 且相互独立，$Y = \max\{X_1, X_2\}$.
- $p_Y(y) = 2(1-e^{-\lambda y})\lambda e^{-\lambda y}$, $y > 0$.

$$
E[\max\{X_1, X_2\}] = \int_0^\infty y 2\lambda(1-e^{-\lambda y})e^{-\lambda y} dy
$$

$$
= \frac{2}{\lambda}\int_0^\infty \lambda y e^{-\lambda y} d(\lambda y) - \frac{1}{2\lambda}\int_0^\infty 2\lambda y e^{-2\lambda y} d(2\lambda y) = (\frac{2}{\lambda} - \frac{1}{2\lambda})\Gamma(2)
$$

# 数学期望和方差的性质

- $E(X+Y) = E(X) + E(Y)$:

- $E(X+Y)=E(X)+E(Y)$:

$$\begin{aligned}E(X+Y)&=\iint_{\mathbb{R}^2}(x+y)p(x,y)\,dxdy\\&=\int_{-\infty}^{\infty}x\int_{-\infty}^{\infty}p(x,y)\,dydx+\int_{-\infty}^{\infty}y\int_{-\infty}^{\infty}p(x,y)\,dxdy\\&=\int_{-\infty}^{\infty}xp_X(x)\,dx+\int_{-\infty}^{\infty}yp_Y(y)\,dy=E(X)+E(Y).\end{aligned}$$

# 数学期望和方差的性质

- $E(X+Y)=E(X)+E(Y)$:

$$
\begin{aligned}
E(X+Y) &= \iint_{\mathbb{R}^2}(x+y)p(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p(x,y)\,dydx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} p(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty} xp_X(x)\,dx + \int_{-\infty}^{\infty} yp_Y(y)\,dy = E(X)+E(Y).
\end{aligned}
$$

- 若 $X$ 与 $Y$ 相互独立，则 $E(XY)=E(X)E(Y)$.

- $E(X+Y) = E(X) + E(Y)$:

$$
\begin{aligned}
E(X+Y) &= \iint_{\mathbb{R}^2} (x+y)p(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p(x,y)\,dydx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} p(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty} xp_X(x)\,dx + \int_{-\infty}^{\infty} yp_Y(y)\,dy = E(X) + E(Y).
\end{aligned}
$$

- 若 $X$ 与 $Y$ 相互独立，则 $E(XY) = E(X)E(Y)$.

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyp_X(x)p_Y(y)\,dxdy \\
&= \int_{-\infty}^{\infty} xp_X(x)\,dx \int_{-\infty}^{\infty} yp_Y(y)\,dy = E(X)E(Y).
\end{aligned}
$$

- 若 $X$ 与 $Y$ 相互独立，则 $Var(X+Y) = Var(X) + Var(Y)$.

# 数学期望和方差的性质

- 若 $X$ 与 $Y$ 相互独立，则 $Var(X+Y) = Var(X) + Var(Y)$.

$$
\begin{aligned}
&Var(X+Y) \\
&= E[(X+Y-E(X+Y))^2] = E[(X-E(X)+Y-E(Y))^2] \\
&= E[(X-E(X))^2] + E[(Y-E(Y))^2] + 2E[(X-E(X))(Y-E(Y) \\
&= Var(X) + Var(Y).
\end{aligned}
$$

# 数学期望和方差的性质

- 若 $X$ 与 $Y$ 相互独立，则 $Var(X+Y) = Var(X) + Var(Y)$.

$$
\begin{aligned}
&Var(X+Y)\\
&= E[(X+Y-E(X+Y))^2] = E[(X-E(X)+Y-E(Y))^2]\\
&= E[(X-E(X))^2] + E[(Y-E(Y))^2] + 2E[(X-E(X))(Y-E(Y))]\\
&= Var(X) + Var(Y).
\end{aligned}
$$

- 前面的三个等式对任意有限个随机变量都成立：
  - $E(X_1+\cdots+X_n) = E(X_1)+\cdots+E(X_n)$,
  - 若 $X_i, i=1,\ldots,n$ 相互独立，则

$$E(X_1\cdots X_n) = E(X_1)\cdots E(X_n),$$

$$Var(X_1+\cdots+X_n) = Var(X_1)+\cdots+Var(X_n).$$

# 例子: 又是抛硬币

- 连续抛硬币一百万次，连续 6 个正面接着连续 6 个反面出现的次数的数学期望值是?

# 例子: 又是抛硬币

- 连续抛硬币一百万次，连续 6 个正面接着连续 6 个反面出现的次数的数学期望值是?
- 考虑随机变量 $I_j$, $I_j = 1$ 如果从第 j 次开始（包括 j）连续抛出了 6 个正面紧接着连续 6 个反面，否则为 0.

# 例子: 又是抛硬币

- 连续抛硬币一百万次，连续 6 个正面接着连续 6 个反面出现的次数的数学期望值是?
- 考虑随机变量 $I_j$, $I_j = 1$ 如果从第 j 次开始（包括 j）连续抛出了 6 个正面紧接着连续 6 个反面，否则为 0.
- 那么出现的次数为 $\sum_{j=1}^{1000000-11} I_j$,

# 例子: 又是抛硬币

- 连续抛硬币一百万次，连续 6 个正面接着连续 6 个反面出现的次数的数学期望值是?
- 考虑随机变量 $I_j$, $I_j = 1$ 如果从第 j 次开始（包括 j）连续抛出了 6 个正面紧接着连续 6 个反面，否则为 0.
- 那么出现的次数为 $\sum_{j=1}^{1000000-11} I_j$,
- $E(I_j) = \frac{1}{2^{12}}$.?
- $E(\sum_{j=1}^{1000000-11} I_j) = \sum_{j=1}^{1000000-11} E(I_j) = \frac{1000000-11}{2^{12}} = 244.14.$

# 回到双色球问题

- 假设有 $N$ 个人各买了一注，考虑随机变量 $X_i$, $X_i = 1$ 若第 i 个人中了大奖，否则 $X_i = 0$.

# 回到双色球问题

- 假设有 $N$ 个人各买了一注，考虑随机变量 $X_i$, $X_i = 1$ 若第 i 个人中了大奖，否则 $X_i = 0$.
- 中大奖的人数为 $Y_N = \sum_{i=1}^{N} X_i$, $E(Y_N) = Np$, $Var(Y_N) = Np(1-p)$.

## 回到双色球问题

- 假设有 $N$ 个人各买了一注，考虑随机变量 $X_i$, $X_i = 1$ 若第 i 个人中了大奖，否则 $X_i = 0$.
- 中大奖的人数为 $Y_N = \sum_{i=1}^{N} X_i$, $E(Y_N) = Np$, $Var(Y_N) = Np(1-p)$.
- 一个中大奖的人都没有的概率是

$$P(Y_N = 0) \leqslant P(|Y - E(Y)| \geqslant Np) \leqslant \frac{Var(Y)}{N^2p^2} = \frac{N(1-p)p}{N^2p^2}.$$

- $p = \frac{1}{17721088}$, $N = 300000000$, $P(Y_N = 0) \leqslant 0.059$.

# 回到双色球问题

- 假设有 $N$ 个人各买了一注，考虑随机变量 $X_i$, $X_i = 1$ 若第 i 个人中了大奖，否则 $X_i = 0$.
- 中大奖的人数为 $Y_N = \sum_{i=1}^{N} X_i$, $E(Y_N) = Np$, $Var(Y_N) = Np(1-p)$.
- 一个中大奖的人都没有的概率是

$$P(Y_N = 0) \leqslant P(|Y - E(Y)| \geqslant Np) \leqslant \frac{Var(Y)}{N^2 p^2} = \frac{N(1-p)p}{N^2 p^2}.$$

- $p = \frac{1}{17721088}$, $N = 300000000$, $P(Y_N = 0) \leqslant 0.059$.更精确一点的计算?

# 回到双色球问题

- 假设有 $N$ 个人各买了一注，考虑随机变量 $X_i$, $X_i = 1$ 若第 i 个人中了大奖，否则 $X_i = 0$.
- 中大奖的人数为 $Y_N = \sum_{i=1}^{N} X_i$, $E(Y_N) = Np$, $Var(Y_N) = Np(1-p)$.
- 一个中大奖的人都没有的概率是

$$P(Y_N = 0) \leqslant P(|Y - E(Y)| \geqslant Np) \leqslant \frac{Var(Y)}{N^2 p^2} = \frac{N(1-p)p}{N^2 p^2}.$$

- $p = \frac{1}{17721088}$, $N = 300000000$, $P(Y_N = 0) \leqslant 0.059$.更精确一点的计算？
- $P(Y_N = 0) = (1-p)^{30000000} = 4.44463 * 10^{-8}$.

# 大数定律：抢先版

- 考虑 $n$ 个独立同分布随机变量, $X_1, \ldots, X_n$, （如 $n$ 次独立重复试验），其方差 $\sigma^2$ 和期望 $\mu$ 均存在（有限）.

# 大数定律：抢先版

- 考虑 $n$ 个独立同分布随机变量, $X_1, \ldots, X_n$, （如 $n$ 次独立重复试验），其方差 $\sigma^2$ 和期望 $\mu$ 均存在（有限）.
- $E(\frac{1}{n}(X_1 + \cdots + X_n)) = \mu$.
- $Var(\frac{1}{n}(X_1 + \cdots + X_n)) = \frac{\sigma^2}{n}$.

# 大数定律：抢先版

- 考虑 $n$ 个独立同分布随机变量, $X_1, \ldots, X_n$, （如 $n$ 次独立重复试验），其方差 $\sigma^2$ 和期望 $\mu$ 均存在（有限）.
- $E(\frac{1}{n}(X_1 + \cdots + X_n)) = \mu$.
- $Var(\frac{1}{n}(X_1 + \cdots + X_n)) = \frac{\sigma^2}{n}$.
- 对于 $\forall \epsilon > 0$,

$$P(|\frac{1}{n}\sum_{i=1}^{n} X_n - \mu| \geqslant \epsilon) \leqslant \frac{\sigma^2}{n\epsilon^2} \to 0, \quad n \to \infty.$$

# 大数定律：抢先版

- 考虑 $n$ 个独立同分布随机变量, $X_1, \ldots, X_n$, （如 $n$ 次独立重复试验）, 其方差 $\sigma^2$ 和期望 $\mu$ 均存在（有限）.
- $E(\frac{1}{n}(X_1 + \cdots + X_n)) = \mu$.
- $Var(\frac{1}{n}(X_1 + \cdots + X_n)) = \frac{\sigma^2}{n}$.
- 对于 $\forall \epsilon > 0$,

$$P(|\frac{1}{n}\sum_{i=1}^{n} X_n - \mu| \geqslant \epsilon) \leqslant \frac{\sigma^2}{n\epsilon^2} \to 0, \quad n \to \infty.$$

- $\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu, \quad n \to \infty.$

- $(X, Y)$ 为二维随机变量，则称

$$Cov(X, Y) = E\Big[(X - E(X))(Y - E(Y))\Big],$$

为 $X$ 与 $Y$ 的协方差。

# 协方差

- $(X, Y)$ 为二维随机变量，则称

$$Cov(X, Y) = E\Big[(X - E(X))(Y - E(Y))\Big],$$

为 $X$ 与 $Y$ 的协方差。

- $Cov(X, Y) = Cov(Y, X)$.

# 协方差

- $(X, Y)$ 为二维随机变量，则称

$$Cov(X, Y) = E\Big[(X - E(X))(Y - E(Y))\Big],$$

为 $X$ 与 $Y$ 的协方差。
- $Cov(X, Y) = Cov(Y, X)$.
- $Cov(X, Y) > 0$, 称 $X$ 与 $Y$ 正相关。

# 协方差

- $(X, Y)$ 为二维随机变量，则称

$$Cov(X, Y) = E\Big[(X - E(X))(Y - E(Y))\Big],$$

为 $X$ 与 $Y$ 的协方差。

- $Cov(X, Y) = Cov(Y, X)$.
- $Cov(X, Y) > 0$, 称 $X$ 与 $Y$ 正相关。
- $Cov(X, Y) < 0$, 称 $X$ 与 $Y$ 负相关。

# 协方差

- $(X, Y)$ 为二维随机变量，则称

$$Cov(X, Y) = E\Big[(X - E(X))(Y - E(Y))\Big],$$

为 $X$ 与 $Y$ 的协方差。
- $Cov(X, Y) = Cov(Y, X)$.
- $Cov(X, Y) > 0$, 称 $X$ 与 $Y$ 正相关。
- $Cov(X, Y) < 0$, 称 $X$ 与 $Y$ 负相关。
- $Cov(X, Y) = 0$, 称 $X$ 与 $Y$ 不相关。

# 协方差的性质

- $Cov(X, Y) = E(XY) - E(X)E(Y)$.

# 协方差的性质

- $Cov(X, Y) = E(XY) - E(X)E(Y)$.

$$
\begin{aligned}
Cov(X, Y) &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\
&= E(XY) - E(X)E(Y).
\end{aligned}
$$

# 协方差的性质

- $Cov(X, Y) = E(XY) - E(X)E(Y)$.

$$\begin{aligned} Cov(X, Y) &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

若 $Cov(X, Y) = 0$, $E(XY) = E(X)E(Y)$.

# 协方差的性质

- $Cov(X,Y)=E(XY)-E(X)E(Y)$.

$$\begin{aligned} Cov(X,Y) &= E[XY-XE(Y)-YE(X)+E(X)E(Y)] \\ &= E(XY)-E(X)E(Y). \end{aligned}$$

  若 $Cov(X,Y)=0$, $E(XY)=E(X)E(Y)$.
- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X,Y)=0$, 反之不然。

# 协方差的性质

- $Cov(X,Y) = E(XY) - E(X)E(Y)$.

$$\begin{aligned} Cov(X,Y) &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

  若 $Cov(X,Y)=0$, $E(XY)=E(X)E(Y)$.
- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X,Y)=0$, 反之不然。
  - $X \sim N(0,1)$, $Y = X^2$, $Cov(X,Y) = E(X^3) - E(X)E(Y) = 0$.

# 协方差的性质

- $Cov(X,Y)=E(XY)-E(X)E(Y)$.

$$\begin{aligned}Cov(X,Y)&=E[XY-XE(Y)-YE(X)+E(X)E(Y)]\\&=E(XY)-E(X)E(Y).\end{aligned}$$

若 $Cov(X,Y)=0$, $E(XY)=E(X)E(Y)$.

- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X,Y)=0$，反之不然。
  - $X\sim N(0,1)$, $Y=X^2$, $Cov(X,Y)=E(X^3)-E(X)E(Y)=0$.
  - $X\sim U(0,2\pi)$, $Y=\cos X$, $Z=\sin X$, $Cov(Y,Z)=0$.

# 协方差的性质

- $Cov(X,Y) = E(XY) - E(X)E(Y)$.

$$\begin{aligned} Cov(X,Y) &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

  若 $Cov(X,Y)=0$, $E(XY)=E(X)E(Y)$.
- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X,Y)=0$, 反之不然。
  - $X \sim N(0,1)$, $Y = X^2$, $Cov(X,Y) = E(X^3) - E(X)E(Y) = 0$.
  - $X \sim U(0,2\pi)$, $Y = \cos X$, $Z = \sin X$, $Cov(Y,Z) = 0$.
- $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

# 协方差的性质

- $Cov(X, Y) = E(XY) - E(X)E(Y)$.

$$\begin{aligned} Cov(X, Y) &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

若 $Cov(X, Y) = 0$, $E(XY) = E(X)E(Y)$.

- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X, Y) = 0$, 反之不然。
  - $X \sim N(0,1)$, $Y = X^2$, $Cov(X, Y) = E(X^3) - E(X)E(Y) = 0$.
  - $X \sim U(0, 2\pi)$, $Y = \cos X$, $Z = \sin X$, $Cov(Y, Z) = 0$.
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

$$\begin{aligned} & Var(X + Y) \\ = & E(X + Y - E(X + Y))^2 \\ = & E(X - E(X))^2 + E(Y - E(Y))^2 + 2E(X - E(X))(Y - E(Y)). \end{aligned}$$

# 协方差的性质

- $Cov(X,Y)=E(XY)-E(X)E(Y)$.

$$\begin{aligned} Cov(X,Y) &= E[XY-XE(Y)-YE(X)+E(X)E(Y)] \\ &= E(XY)-E(X)E(Y). \end{aligned}$$

若 $Cov(X,Y)=0$, $E(XY)=E(X)E(Y)$.

- 若 $X$ 与 $Y$ 相互独立，则 $Cov(X,Y)=0$, 反之不然。
  - $X\sim N(0,1)$, $Y=X^2$, $Cov(X,Y)=E(X^3)-E(X)E(Y)=0$.
  - $X\sim U(0,2\pi)$, $Y=\cos X$, $Z=\sin X$, $Cov(Y,Z)=0$.
- $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$

$$\begin{aligned} & Var(X+Y) \\ &= E(X+Y-E(X+Y))^2 \\ &= E(X-E(X))^2+E(Y-E(Y))^2+2E(X-E(X))(Y-E(Y)). \end{aligned}$$

若 $Cov(X,Y)=0$, $Var(X+Y)=Var(X)+Var(Y)$.

# 协方差的性质

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$

# 协方差的性质

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$
- $\forall a, b \in \mathbb{R},\ Cov(aX, bY) = abCov(X, Y).$

# 协方差的性质

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$
- $\forall a, b \in \mathbb{R},\ Cov(aX, bY) = abCov(X, Y).$
- 分配律：$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$。

# 协方差的性质

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$
- $\forall a, b \in \mathbb{R},\ Cov(aX, bY) = abCov(X, Y).$
- 分配律：$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$。

$$\begin{aligned} Cov(X + Y, Z) &= E((X + Y)Z) - E(X + Y)E(Z) \\ &= E(XZ) + E(YZ) - E(X)E(Z) - E(Y)E(Z) \\ &= Cov(X, Z) + Cov(Y, Z). \end{aligned}$$

- 施瓦茨（Schwarz）不等式：$[Cov(X, Y)]^2 \leqslant \sigma_X^2 \sigma_Y^2.$

# 协方差的性质

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$
- $\forall a, b \in \mathbb{R},\ Cov(aX, bY) = abCov(X, Y).$
- 分配律：$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$。

$$\begin{aligned} Cov(X + Y, Z) &= E((X + Y)Z) - E(X + Y)E(Z) \\ &= E(XZ) + E(YZ) - E(X)E(Z) - E(Y)E(Z) \\ &= Cov(X, Z) + Cov(Y, Z). \end{aligned}$$

- 施瓦茨（Schwarz）不等式：$[Cov(X, Y)]^2 \leqslant \sigma_X^2\sigma_Y^2$.考虑

$$g(t) = E[t(X-E(X))+(Y-E(Y))]^2 = t^2\sigma_X^2+2tCov(X, Y)+\sigma_Y^2 \geqslant 0,$$

- $\forall a \in \mathbb{R},\ Cov(X, a) = 0.$
- $\forall a, b \in \mathbb{R},\ Cov(aX, bY) = abCov(X, Y).$
- 分配律：$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$。

$$\begin{aligned} Cov(X+Y, Z) &= E((X+Y)Z) - E(X+Y)E(Z) \\ &= E(XZ) + E(YZ) - E(X)E(Z) - E(Y)E(Z) \\ &= Cov(X, Z) + Cov(Y, Z). \end{aligned}$$

- 施瓦茨（Schwarz）不等式：$[Cov(X, Y)]^2 \leqslant \sigma_X^2\sigma_Y^2$.考虑

$$g(t) = E[t(X-E(X))+(Y-E(Y))]^2 = t^2\sigma_X^2+2tCov(X, Y)+\sigma_Y^2 \geqslant 0,$$

则

$$(2Cov(X, Y))^2 - 4\sigma_X^2\sigma_Y^2 \leqslant 0.$$

## 计算协方差

- 设二维随机变量 $(X, Y)$ 在矩形

$$G = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\},$$

上服从均匀分布，记

$$U = \begin{cases} 1, & X > Y, \\ 0, & X \leqslant Y, \end{cases} \quad V = \begin{cases} 1, & X > 2Y, \\ 0, & X \leqslant 2Y. \end{cases}$$

求 $U$ 和 $V$ 的协方差。

- 设二维随机变量 $(X, Y)$ 在矩形
$$G = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\},$$
上服从均匀分布，记
$$U = \begin{cases} 1, & X > Y, \\ 0, & X \leqslant Y, \end{cases} \quad V = \begin{cases} 1, & X > 2Y, \\ 0, & X \leqslant 2Y. \end{cases}$$
求 $U$ 和 $V$ 的协方差。
- $p(x, y) = \begin{cases} \frac{1}{2}, & (x, y) \in G \\ 0, & \text{其他}. \end{cases}$

## 计算协方差

- 设二维随机变量 $(X, Y)$ 在矩形
$$G = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\},$$
上服从均匀分布，记
$$U = \begin{cases} 1, & X > Y, \\ 0, & X \leqslant Y, \end{cases} \quad V = \begin{cases} 1, & X > 2Y, \\ 0, & X \leqslant 2Y. \end{cases}$$
求 $U$ 和 $V$ 的协方差。
- $p(x, y) = \begin{cases} \frac{1}{2}, & (x, y) \in G \\ 0, & \text{其他}. \end{cases}$
- $P(U = 1) = P(X > 1) + \int_0^1 \int_0^x 0.5dydx = 0.75,$
$P(U = 0) = 0.25,\ P(V = 1) = \int_0^2 \int_0^{x/2} 0.5dydx = 0.5,$
$P(V = 0) = 0.5.$

# 计算协方差

- 设二维随机变量 $(X, Y)$ 在矩形
$$G = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\},$$
上服从均匀分布，记
$$U = \begin{cases} 1, & X > Y, \\ 0, & X \leqslant Y, \end{cases} \quad V = \begin{cases} 1, & X > 2Y, \\ 0, & X \leqslant 2Y. \end{cases}$$
求 $U$ 和 $V$ 的协方差。
- $p(x, y) = \begin{cases} \frac{1}{2}, & (x, y) \in G \\ 0, & \text{其他}. \end{cases}$
- $P(U = 1) = P(X > 1) + \int_0^1 \int_0^x 0.5 dy dx = 0.75,$
$P(U = 0) = 0.25,\ P(V = 1) = \int_0^2 \int_0^{x/2} 0.5 dy dx = 0.5,$
$P(V = 0) = 0.5.$
- $P(UV = 1) = \int_0^2 \int_0^{x/2} 0.5 dy dx = 0.5.$

# 计算协方差

- 设二维随机变量 $(X, Y)$ 在矩形
$$G = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\},$$
上服从均匀分布，记
$$U = \begin{cases} 1, & X > Y, \\ 0, & X \leqslant Y, \end{cases} \quad V = \begin{cases} 1, & X > 2Y, \\ 0, & X \leqslant 2Y. \end{cases}$$
求 $U$ 和 $V$ 的协方差。
- $p(x, y) = \begin{cases} \frac{1}{2}, & (x, y) \in G \\ 0, & \text{其他}. \end{cases}$
- $P(U = 1) = P(X > 1) + \int_0^1 \int_0^x 0.5dydx = 0.75,$
$P(U = 0) = 0.25,\ P(V = 1) = \int_0^2 \int_0^{x/2} 0.5dydx = 0.5,$
$P(V = 0) = 0.5.$
- $P(UV = 1) = \int_0^2 \int_0^{x/2} 0.5dydx = 0.5.$
- $Cov(U, V) = E(UV) - E(U)E(V) = 0.5 - 0.75 * 0.5 = 0.125.$

# 相关系数

- 设 $(X,Y)$ 是一个二维随机变量，且 $Var(X)=\sigma_X^2>0$, $Var(Y)=\sigma_Y^2>0$, 则称

$$Corr(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}=\frac{Cov(X,Y)}{\sigma_X,\sigma_Y},$$

为 $X$ 与 $Y$ 的相关系数。

# 相关系数

- 设 $(X, Y)$ 是一个二维随机变量，且 $Var(X) = \sigma_X^2 > 0$, $Var(Y) = \sigma_Y^2 > 0$, 则称

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X, Y)}{\sigma_X, \sigma_Y},$$

为 $X$ 与 $Y$ 的相关系数。

- 相关系数是相应标准化随机变量的协方差:

## 相关系数

- 设 $(X, Y)$ 是一个二维随机变量，且 $Var(X) = \sigma_X^2 > 0$, $Var(Y) = \sigma_Y^2 > 0$, 则称

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)\,Var(Y)}} = \frac{Cov(X, Y)}{\sigma_X, \sigma_Y},$$

为 $X$ 与 $Y$ 的相关系数。

- 相关系数是相应标准化随机变量的协方差：

$$X' = \frac{X - \mu_X}{\sigma_X}, \quad Y' = \frac{Y - \mu_Y}{\sigma_Y},$$

- 设 $(X, Y)$ 是一个二维随机变量，且 $Var(X) = \sigma_X^2 > 0$, $Var(Y) = \sigma_Y^2 > 0$, 则称

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)\,Var(Y)}} = \frac{Cov(X, Y)}{\sigma_X, \sigma_Y},$$

为 $X$ 与 $Y$ 的相关系数。

- 相关系数是相应标准化随机变量的协方差：

$$X' = \frac{X - \mu_X}{\sigma_X}, \quad Y' = \frac{Y - \mu_Y}{\sigma_Y},$$

$$Cov(X', Y') = Cov(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y}) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}.$$

# 相关系数的性质

- 有界性：$|Corr(X, Y)| \leqslant 1$.

# 相关系数的性质

- 有界性：$|Corr(X, Y)| \leqslant 1$.
- $Corr(X, Y) = \pm 1$ 当且仅当 $X$ 与 $Y$ 几乎处处线性相关，即存在 $a \neq 0, b$, 使得
$$P(X = aY + b) = 1.$$

# 相关系数的性质

- 有界性：$|Corr(X,Y)| \leqslant 1$.
- $Corr(X,Y) = \pm 1$ 当且仅当 $X$ 与 $Y$ 几乎处处线性相关，即存在 $a \neq 0, b$，使得
$$P(X = aY + b) = 1.$$
  - 充分性：若 $Y = aX + b$，则 $Corr(X,Y) = \frac{a\sigma_X^2}{|a|\sigma_X^2} = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$.

# 相关系数的性质

- 有界性：$|Corr(X,Y)| \leqslant 1.$
- $Corr(X,Y) = \pm 1$ 当且仅当 $X$ 与 $Y$ 几乎处处线性相关，即存在 $a \neq 0, b$，使得
$$P(X = aY + b) = 1.$$

  - 充分性：若 $Y = aX + b$，则 $Corr(X,Y) = \frac{a\sigma_X^2}{|a|\sigma_X^2} = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$.
  - 必要性：$Var(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}) = 2[1 \pm Corr(X,Y)].$

# 相关系数的性质

- 有界性：$|Corr(X, Y)| \leqslant 1$.
- $Corr(X, Y) = \pm 1$ 当且仅当 $X$ 与 $Y$ 几乎处处线性相关，即存在 $a \neq 0, b$, 使得
$$P(X = aY + b) = 1.$$

  - 充分性：若 $Y = aX + b$, 则 $Corr(X, Y) = \frac{a\sigma_X^2}{|a|\sigma_X^2} = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$.
  - 必要性：$Var(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}) = 2[1 \pm Corr(X, Y)]$.
    若 $Corr(X, Y) = 1$, 则
$$Var(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}) = 0 \Rightarrow P(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c) = 1.$$

# 相关系数的性质

- 有界性：$|Corr(X,Y)| \leqslant 1$.
- $Corr(X,Y) = \pm 1$ 当且仅当 $X$ 与 $Y$ 几乎处处线性相关，即存在 $a \neq 0, b$，使得
$$P(X = aY + b) = 1.$$

  - 充分性：若 $Y = aX + b$，则 $Corr(X,Y) = \frac{a\sigma_X^2}{|a|\sigma_X^2} = \begin{cases} 1, & a > 0 \\ -1, & a < 0. \end{cases}$.
  - 必要性：$Var(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}) = 2[1 \pm Corr(X,Y)]$.
    若 $Corr(X,Y) = 1$，则
$$Var(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}) = 0 \Rightarrow P(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y} = c) = 1.$$
    若 $Corr(X,Y) = -1$，则
$$Var(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}) = 0 \Rightarrow P(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = c) = 1.$$

# 协方差或者相关系数的一个应用: 投资组合

- 有金融产品 $X$ 和 $Y$，考虑投资组合 $X+Y$.
- 回报期望为 $E(X)+E(Y)$.
- 组合的方差为 $Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$.
- 在回报期望一样的时候，希望风险越小越好。
- 应该选择 $Cov(X,Y)<0$ 的产品。

# 二维正态分布的相关系数

- $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

# 二维正态分布的相关系数

- $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
-

$$
\begin{aligned}
Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_1)(y-\mu_2) \\
&\times \quad \exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \\
&\quad + \frac{(y-\mu_2)^2}{\sigma_2^2}]\} dxdy
\end{aligned}
$$

# 二维正态分布的相关系数

- $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
-

$$
\begin{aligned}
Cov(X, Y) = & E[(X - E(X))(Y - E(Y))] \\
= & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_1)(y-\mu_2) \\
\times & \quad \exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \\
& + \frac{(y-\mu_2)^2}{\sigma_2^2}]\} dxdy
\end{aligned}
$$

将中括号内化为

$$
(\frac{x-\mu}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2})^2 + (1-\rho^2)(\frac{y-\mu_2}{\sigma_2})^2.
$$

# 二维正态分布的相关系数

- 作变量代换

$$\begin{cases} u = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}\right) \\ v = \frac{y-\mu_2}{\sigma_2}, \end{cases} \quad 即 \quad \begin{cases} x - \mu_1 = \sigma_1(u\sqrt{1-\rho^2} + \rho v), \\ y - \mu_2 = \sigma_2 v, \end{cases}$$

# 二维正态分布的相关系数

- 作变量代换

$$\begin{cases} u = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}\right) \\ v = \frac{y-\mu_2}{\sigma_2}, \end{cases} \quad \text{即} \quad \begin{cases} x-\mu_1 = \sigma_1(u\sqrt{1-\rho^2}+\rho v), \\ y-\mu_2 = \sigma_2 v, \end{cases}$$

- $|J| = \sigma_1\sigma_2\sqrt{1-\rho^2}$.

# 二维正态分布的相关系数

- 作变量代换

$$\begin{cases} u = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}\right) \\ v = \frac{y-\mu_2}{\sigma_2}, \end{cases} \quad \text{即} \quad \begin{cases} x-\mu_1 = \sigma_1(u\sqrt{1-\rho^2}+\rho v), \\ y-\mu_2 = \sigma_2 v, \end{cases}$$

- $|J| = \sigma_1\sigma_2\sqrt{1-\rho^2}$.
- 

$$\begin{aligned} & Cov(X,Y) \\ &= \frac{\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(uv\sqrt{1-\rho^2}+\rho v^2)\exp\{-\frac{1}{2}(u^2+v^2)\}dudv \\ &= \rho\sigma_1\sigma_2. \end{aligned}$$

# 二维正态分布的相关系数

- 作变量代换

$$\begin{cases} u = \frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}\right) \\ v = \frac{y-\mu_2}{\sigma_2}, \end{cases} \quad \text{即} \quad \begin{cases} x-\mu_1 = \sigma_1(u\sqrt{1-\rho^2}+\rho v), \\ y-\mu_2 = \sigma_2 v, \end{cases}$$

- $|J| = \sigma_1\sigma_2\sqrt{1-\rho^2}$.
-

$$\begin{aligned} & Cov(X,Y) \\ &= \frac{\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(uv\sqrt{1-\rho^2}+\rho v^2)\exp\{-\frac{1}{2}(u^2+v^2)\}dudv \\ &= \rho\sigma_1\sigma_2. \end{aligned}$$

- $Corr(X,Y) = \rho$.

# 二维正态分布的相关系数

- 作变量代换

$$\begin{cases} u = \frac{1}{\sqrt{1-\rho^2}}(\frac{x-\mu_1}{\sigma_1} - \rho\frac{y-\mu_2}{\sigma_2}) \\ v = \frac{y-\mu_2}{\sigma_2}, \end{cases} \quad \text{即} \quad \begin{cases} x-\mu_1 = \sigma_1(u\sqrt{1-\rho^2}+\rho v), \\ y-\mu_2 = \sigma_2 v, \end{cases}$$

- $|J| = \sigma_1\sigma_2\sqrt{1-\rho^2}$.
-

$$\begin{aligned} & Cov(X, Y) \\ & = \frac{\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(uv\sqrt{1-\rho^2}+\rho v^2)\exp\{-\frac{1}{2}(u^2+v^2)\}dudv \\ & = \rho\sigma_1\sigma_2. \end{aligned}$$

- $Corr(X, Y) = \rho$.
- 二维正态分布 $(X, Y)$：$X$, $Y$ 不相关与相互独立等价。（非常特殊）

- 若 $n$ 维随机变量 $X=(X_1,\ldots,X_n)$ 的每个分量的数学期望都存在，则 $n$ 维随机变量的数学期望为：

$$E(X)=(E(X_1),\ldots,E(X_n)),$$

- 其协方差矩阵为：

$$\begin{aligned}&E[(X-E(X))(X-E(X))']\\&=\begin{pmatrix}Var(X_1) & Cov(X_1,X_2) & \cdots & Cov(X_1,X_n)\\ Cov(X_2,X_1) & Var(X_2) & \cdots & Cov(X_2,X_n)\\ \vdots & \vdots & & \vdots\\ Cov(X_n,X_1) & Cov(X_n,X_2) & \cdots & Var(X_n)\end{pmatrix}.\end{aligned}$$

- 若 $n$ 维随机变量 $X=(X_1,\ldots,X_n)$ 的每个分量的数学期望都存在，则 $n$ 维随机变量的数学期望为：

$$E(X)=(E(X_1),\ldots,E(X_n)),$$

- 其协方差矩阵为：

$$\begin{aligned}&E[(X-E(X))(X-E(X))']\\&=\begin{pmatrix}Var(X_1) & Cov(X_1,X_2) & \cdots & Cov(X_1,X_n)\\ Cov(X_2,X_1) & Var(X_2) & \cdots & Cov(X_2,X_n)\\ \vdots & \vdots & & \vdots\\ Cov(X_n,X_1) & Cov(X_n,X_2) & \cdots & Var(X_n)\end{pmatrix}.\end{aligned}$$

- 协方差矩阵为非负定对称矩阵。

# $n$ 维正态分布

- $n$ 维正态分布的密度函数有以下形式:

$$p(x_1, \ldots, x_n) = p(x) = \frac{1}{(2\pi)^{n/2}|B|^{1/2}} \exp\{-\frac{1}{2}(x - a)' B^{-1}(x - a)\},$$

其中 $B$ 是其协方差矩阵，$a$ 为其数学期望（向量）。

# $n$ 维正态分布

- $n$ 维正态分布的密度函数有以下形式：

$$p(x_1, \ldots, x_n) = p(x) = \frac{1}{(2\pi)^{n/2}|B|^{1/2}} \exp\{-\frac{1}{2}(x-a)'B^{-1}(x-a)\},$$

其中 $B$ 是其协方差矩阵，$a$ 为其数学期望（向量）。

- 二维时

$$a = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad B = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

# 离散随机变量的条件分布

- 二维离散随机变量 $(X, Y)$ 的联合分布列为

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j = 1, 2, \ldots .$$

- 对一切使得 $P(Y = y_j) = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij} > 0$ 的 $y_j$，称

$$p_{i|j} = P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \ldots ,$$

为给定 $Y = y_j$ 条件下 $X$ 的条件分布列。

# 离散随机变量的条件分布

- 二维离散随机变量 $(X, Y)$ 的联合分布列为

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j = 1, 2, \ldots.$$

- 对一切使得 $P(Y = y_j) = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij} > 0$ 的 $y_j$，称

$$p_{i|j} = P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \ldots,$$

为给定 $Y = y_j$ 条件下 $X$ 的条件分布列。为什么是分布列?

- 二维离散随机变量 $(X, Y)$ 的联合分布列为

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j = 1, 2, \ldots.$$

- 对一切使得 $P(Y = y_j) = p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij} > 0$ 的 $y_j$，称

$$p_{i|j} = P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \ldots,$$

为给定 $Y = y_j$ 条件下 $X$ 的条件分布列。为什么是分布列?

- 给定 $Y = y_j$ 条件下 $X$ 的条件分布函数为

$$F(x|y_j) = \sum_{x_i \leqslant x} P(X = x_i | Y = y_j) = \sum_{x_i \leqslant x} p_{i|j}.$$

- 随机变量 $X$ 与 $Y$ 相互独立，且 $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$. 求 $X + Y = n$ 条件下 $X$ 的条件分布。

# 例子

- 随机变量 $X$ 与 $Y$ 相互独立，且 $X \sim P(\lambda_1)$，$Y \sim P(\lambda_2)$. 求 $X+Y=n$ 条件下 $X$ 的条件分布。
- $X+Y \sim P(\lambda_1+\lambda_2)$.
-

$$
\begin{aligned}
P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} \\
&= \frac{P(X=k)P(X=n-k)}{P(X+Y=n)} \\
&= \frac{\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}}{\frac{(\lambda_1+\lambda_2)^n}{n!}e^{-\lambda_1-\lambda_2}} \\
&= \frac{n!}{k!(n-k)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.
\end{aligned}
$$

# 例子

- 随机变量 $X$ 与 $Y$ 相互独立，且 $X \sim P(\lambda_1)$，$Y \sim P(\lambda_2)$. 求 $X+Y=n$ 条件下 $X$ 的条件分布。
- $X+Y \sim P(\lambda_1+\lambda_2)$.
- $$\begin{aligned} P(X=k|X+Y=n) &= \frac{P(X=k, X+Y=n)}{P(X+Y=n)} \\ &= \frac{P(X=k)P(X=n-k)}{P(X+Y=n)} \\ &= \frac{\frac{\lambda_1^k}{k!}e^{-\lambda_1}\frac{\lambda_2^{n-k}}{(n-k)!}e^{-\lambda_2}}{\frac{(\lambda_1+\lambda_2)^n}{n!}e^{-\lambda_1-\lambda_2}} \\ &= \frac{n!}{k!(n-k)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}. \end{aligned}$$
- 这是二项分布。

- 二维随机变量 $(X, Y)$ 的联合密度函数为 $p(x, y)$，边际密度函数分别为 $p_X(x)$，$p_Y(y)$. 对于使得 $p_Y(y) > 0$ 的 $y$，在给定 $Y = y$ 的条件下 $X$ 的条件分布函数和条件密度函数分别为

$$F(x|y) = \int_{-\infty}^{x} \frac{p(u, y)}{p_Y(y)} du, \quad p(x|y) = \frac{p(x, y)}{p_Y(y)}.$$

- 二维随机变量 $(X, Y)$ 的联合密度函数为 $p(x, y)$, 边际密度函数分别为 $p_X(x)$, $p_Y(y)$. 对于使得 $p_Y(y) > 0$ 的 $y$, 在给定 $Y = y$ 的条件下 $X$ 的条件分布函数和条件密度函数分别为

$$F(x|y) = \int_{-\infty}^{x} \frac{p(u, y)}{p_Y(y)} du, \quad p(x|y) = \frac{p(x, y)}{p_Y(y)}.$$

$$\begin{aligned}
P(X \leqslant x | Y = y) &= \lim_{h \to 0} P(X \leqslant x | y \leqslant Y \leqslant y + h) \\
&= \lim_{h \to 0} \frac{P(X \leqslant x, y \leqslant Y \leqslant y + h)}{P(y \leqslant Y \leqslant y + h)} \\
&= \lim_{h \to 0} \frac{\int_{-\infty}^{x} \int_{y}^{y+h} p(x', y') dy' dx'}{\int_{y}^{y+h} p_Y(y') dy'} \\
&= \int_{-\infty}^{x} \frac{p(x', y)}{p_Y(y)} dx'.
\end{aligned}$$

# 例子

- $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

# 例子

- $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.
- $X$ 在 $Y = y$ 条件下的条件密度函数为

$$
\begin{aligned}
& p(x|y) \\
&= \frac{p(x, y)}{p_Y(y)} \\
&= \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\{-\frac{1}{2(1-\rho^2)}(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2})\}}{\frac{1}{\sqrt{2\pi}\sigma_2} \exp\{-\frac{(y-\mu_2)}{2\sigma_2^2}\}} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} \exp\{-\frac{1}{2\sigma_1^2(1-\rho^2)}[x - (\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2))]^2\} \\
&\sim N(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2)).
\end{aligned}
$$

# 连续场合的全概率公式和贝叶斯公式

- $p(x,y) = p_X(x)p(y|x) = p_Y(y)p(x|y)$.

# 连续场合的全概率公式和贝叶斯公式

- $p(x,y) = p_X(x)p(y|x) = p_Y(y)p(x|y)$.
- 全概率公式:

$$p_Y(y) = \int_{-\infty}^{\infty} p_X(x)p(y|x)\,dx,$$

$$p_X(x) = \int_{-\infty}^{\infty} p_Y(y)p(x|y)\,dy.$$

# 连续场合的全概率公式和贝叶斯公式

- $p(x, y) = p_X(x)p(y|x) = p_Y(y)p(x|y)$.
- 全概率公式:

$$p_Y(y) = \int_{-\infty}^{\infty} p_X(x)p(y|x)\,dx,$$

$$p_X(x) = \int_{-\infty}^{\infty} p_Y(y)p(x|y)\,dy.$$

- 贝叶斯公式:

$$p(x|y) = \frac{p_X(x)p(y|x)}{\int_{-\infty}^{\infty} p_X(x)p(y|x)\,dx}.$$

$$p(y|x) = \frac{p_Y(y)p(x|y)}{\int_{-\infty}^{\infty} p_Y(y)p(x|y)\,dy}.$$

# 条件期望

- 条件分布的数学期望称为条件期望，即

$$E(X|Y=y)=\begin{cases}\sum_i x_i P(X=x_i|Y=y), & \text{离散情形}\\ \int_{-\infty}^{\infty} xp(x|y)\,dx, & \text{连续情形}.\end{cases}$$

$$E(Y|X=x)=\begin{cases}\sum_j y_j P(Y=y_j|X=x), & \text{离散情形}\\ \int_{-\infty}^{\infty} yp(y|x)\,dy, & \text{连续情形}.\end{cases}$$

# 条件期望

- 条件分布的数学期望称为条件期望，即

$$E(X|Y=y)=\begin{cases}\sum_i x_i P(X=x_i|Y=y), & \text{离散情形}\\ \int_{-\infty}^{\infty} xp(x|y)\,dx, & \text{连续情形.}\end{cases}$$

$$E(Y|X=x)=\begin{cases}\sum_j y_j P(Y=y_j|X=x), & \text{离散情形}\\ \int_{-\infty}^{\infty} yp(y|x)\,dy, & \text{连续情形.}\end{cases}$$

- 条件期望 $E(X|Y=y)$ 是 $y$ 的函数，若令 $g(y)=E(X|Y=y)$，则 $g(Y)=E(X|Y)$ 是一个随机变量，我们称为 $X$ 关于 $Y$ 的条件期望。

# 条件期望的性质

- 重期望公式 $E(X) = E(E(X|Y))$.

# 条件期望的性质

- 重期望公式 $E(X) = E(E(X|Y))$.

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xp(x,y)\,dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xp(x|y)p_Y(y)\,dydx \\
&= \int_{-\infty}^{\infty}\Big[\int_{-\infty}^{\infty} xp(x|y)\,dx\Big]p_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} E(X|Y=y)p_Y(y)\,dy = E(E(X|Y)).
\end{aligned}
$$

- 重期望公式 $E(X) = E(E(X|Y))$.

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xp(x,y)\,dydx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xp(x|y)p_Y(y)\,dydx \\
&= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} xp(x|y)\,dx\right]p_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} E(X|Y=y)p_Y(y)\,dy = E(E(X|Y)).
\end{aligned}
$$

# 条件期望的性质

- 如果 $X$ 与 $Y$ 相互独立，则 $E(X|Y) = E(X)$.

$$E(X|Y=y) = \int xp(x|y)\,dx = \int xp(x)\,dx = E(X).$$

# 条件期望的性质

- 如果 $X$ 与 $Y$ 相互独立，则 $E(X|Y) = E(X)$.

$$E(X|Y=y) = \int xp(x|y)\,dx = \int xp(x)\,dx = E(X).$$

- $E(f(Y)Z|Y) = f(Y)E(Z|Y)$.

# 条件期望的性质

- 如果 $X$ 与 $Y$ 相互独立，则 $E(X|Y) = E(X)$.

$$E(X|Y=y) = \int xp(x|y)\,dx = \int xp(x)\,dx = E(X).$$

- $E(f(Y)Z|Y) = f(Y)E(Z|Y)$.

$$\begin{aligned} E(f(Y)Z|Y=y) &= \int_{-\infty}^{\infty} f(y)zp(f(y)z|y)\,dz \\ &= f(y)\int_{-\infty}^{\infty} zp(z|y)\,dz = f(y)E(Z|Y=y). \end{aligned}$$

# 随机个随机变量和的数学期望

- 设 $X_1, \ldots,$ 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 $\{X_1, \ldots\}$ 独立，则

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

# 随机个随机变量和的数学期望

- 设 $X_1, \ldots,$ 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 $\{X_1, \ldots\}$ 独立，则

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

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