概率统计第八讲：函数分布

史灵生　清华数学系

1、本讲提要

2、光滑双射密度公式

定理

若随机变量(X,Y)的联合密度函数为$p(x,y)$，则光滑双射$(U,V)=G(X,Y)$给出的随机变量(U,V)的联合密度为

$$p_{U,V}(u,v)=p(x(u,v),y(u,v))|J(u,v)|,\ (u,v)\in R(G),$$

$$J(u,v)=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}=\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}-\frac{\partial x}{\partial v}\frac{\partial y}{\partial u}.$$

$$\begin{aligned}P[(U,V)\in A] &= P[(X,Y)\in G^{-1}A]=\iint_{G^{-1}A}p(x,y)dxdy\\ &= \iint_A p(x(u,v),y(u,v))|J(u,v)|dudv.\end{aligned}$$

3、线性映射的分布

推论

设$\det A\neq 0$且$Y=AX+b$，则$p_Y=p_X(A^{-1}Y-A^{-1}b)/|\det A|$。

正态分布在线性变换下的不变性

若$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，令$X^*=\frac{X-\mu_1}{\sigma_1}$，$Y^*=\frac{Y-\mu_2}{\sigma_2}$，则$(X^*,Y^*)\sim N(0,0,1,1,\rho)$。

- $p_{X,Y}(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{Q(x,y)}{2}\right)$,
- $\binom{X^*}{Y^*}=\begin{pmatrix}1/\sigma_1 & 0\\ 0 & 1/\sigma_2\end{pmatrix}\binom{X}{Y}-\binom{\mu_1/\sigma_1}{\mu_2/\sigma_2}$,
- $\binom{X}{Y}=\begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix}\binom{X^*}{Y^*}+\binom{\mu_1}{\mu_2}$,
- $p_{X^*,Y^*}(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}$。

4、和的分布

定理

若(X,Y)的联合密度函数为$p(x,y)$，则$U=X+Y$的密度函数为
$p_U(u)=\int_{-\infty}^{\infty}p(u-v,v)dv$。

证明：

- 令$V=Y$，
- 则$\binom{U}{V}=\begin{pmatrix}1&1\\0&1\end{pmatrix}\binom{X}{Y}$，
- $\binom{X}{Y}=\begin{pmatrix}1&1\\0&1\end{pmatrix}^{-1}\binom{U}{V}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}\binom{U}{V}$，
- $p_{U,V}(u,v)=p(u-v,v)\left|\det\begin{pmatrix}1&1\\0&1\end{pmatrix}\right|^{-1}=p(u-v,v)$，
- $p_U(u)=\int_{-\infty}^{\infty}p_{U,V}(u,v)dv=\int_{-\infty}^{\infty}p(u-v,v)dv$。

5、积的分布

例

(X,Y)的联合密度函数为$p(x,y)$，则$U=XY$的密度函数为
$p_U(u)=\int_{-\infty}^{\infty}p(u/v,v)/|v|dv$。

- 令$V=Y$，则$\binom{U}{V}=\binom{XY}{Y}$。
- $\binom{X}{Y}=\binom{U/V}{V}$，
- $J=\begin{vmatrix}1/v&-u/v^2\\0&1\end{vmatrix}=1/v$，
- $p_{U,V}(u,v)=p(u/v,v)|J|=p(u/v,v)/|v|$，
- $p_U(u)=\int_{-\infty}^{\infty}p_{U,V}(u,v)dv=\int_{-\infty}^{\infty}p(u/v,v)/|v|dv$。

推论

若X,Y相互独立，则$U=XY$的密度函数为
$$p_U(u)=\int_{-\infty}^{\infty}p_X(u/v)p_Y(v)/|v|dv。$$

6、商的分布

例

(X,Y)的联合密度函数为$p(x,y)$，则$U=X/Y$的密度函数为
$p_U(u)=\int_{-\infty}^{\infty}p(uv,v)|v|dv$。

- 令$V=Y$，则$\binom{U}{V}=\binom{X/Y}{Y}$。
- $\binom{X}{Y}=\binom{UV}{V}$，
- $J=\begin{vmatrix}v&u\\0&1\end{vmatrix}=v$，
- $p_{U,V}(u,v)=p(uv,v)|J|=p(uv,v)|v|$，
- $p_U(u)=\int_{-\infty}^{\infty}p_{U,V}(u,v)dv=\int_{-\infty}^{\infty}p(uv,v)|v|dv$。

推论

若X,Y相互独立，则$U=X/Y$的密度函数为
$$p_U(u)=\int_{-\infty}^{\infty}p_X(uv)p_Y(v)|v|dv。$$

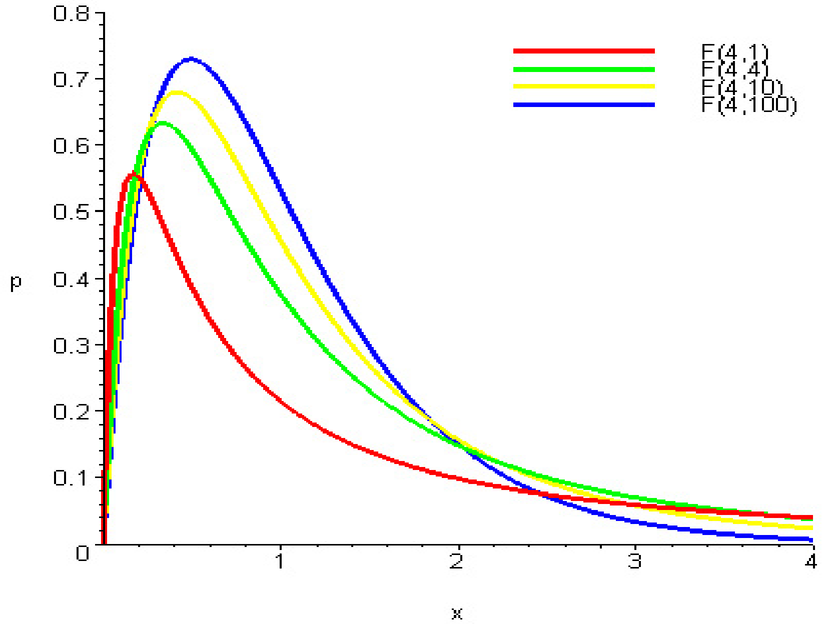
7、F分布

定义

设$X\sim\chi^2(m)$，$Y\sim\chi^2(n)$相互独立，则称$F=\frac{X/m}{Y/n}$的分布是自由度为m与n的F分布，记为$F\sim F(m,n)$，其中m为分子自由度，n为分母自由度。

$$\begin{aligned}
p_{\frac{X}{Y}}(z) &= \int_0^{\infty}yp_X(zy)p_Y(y)dy\\
&= \frac{z^{m/2-1}}{\Gamma(m/2)\Gamma(n/2)2^{(m+n)/2}}\int_0^{\infty}y^{\frac{m+n}{2}-1}e^{\frac{-y(1+z)}{2}}dy\\
&\overset{u=\frac{y(1+z)}{2}}{=} \frac{z^{m/2-1}(1+z)^{-(m+n)/2}}{\Gamma(m/2)\Gamma(n/2)}\int_0^{\infty}u^{(m+n)/2-1}e^{-u}du\\
&= \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)}z^{m/2-1}(1+z)^{-(m+n)/2},\ z>0.\\
p_F(x) &= \frac{m}{n}p_{\frac{X}{Y}}\left(\frac{mx}{n}\right)=\frac{(\frac{m}{n})^{\frac{m}{2}}\Gamma(\frac{m+n}{2})}{\Gamma(m/2)\Gamma(n/2)}x^{\frac{m}{2}-1}\left(1+\frac{mx}{n}\right)^{-\frac{m+n}{2}}.
\end{aligned}$$

8、F分布



9、t分布

定义

设$X \sim N(0,1)$，$Y \sim \chi^2(n)$相互独立，则称$t = \frac{X}{\sqrt{Y/n}}$的分布是自由度为n的t分布，记为$t \sim t(n)$。

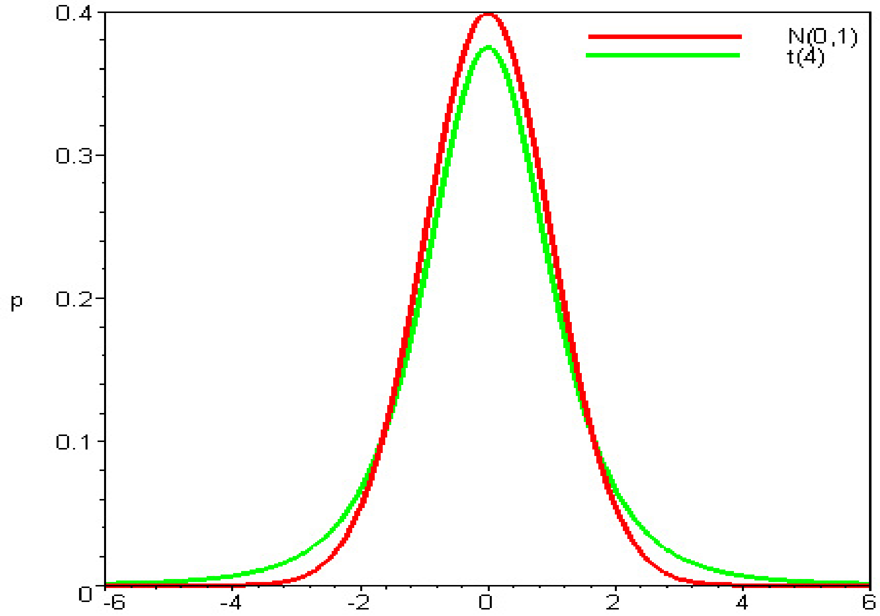
t分布的密度：

- 由X与$-X$同分布知t与$-t$同分布。
- 不失一般性，设$x > 0$，则

$F_t(x) - F_t(0) = P(0 < t \le x) = P(-x \le t < 0) = \frac{1}{2}P(t^2 \le x^2)$，

- 于是$F_t(x) = F_t(0) + \frac{1}{2}P(t^2 \le x^2) = \frac{1}{2} + \frac{1}{2}F_{t^2}(x^2)$。
- 又$t^2 = \frac{X^2}{Y/n} \sim F(1,n)$，
- 故$p_t(x) = xp_{t^2}(x^2) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1 + x^2/n)^{-(n+1)/2}$。

10、t分布



11、函数的数学期望

定理

- 若$g : \mathbb{R}^2 \to \mathbb{R}$是二元(Borel)函数且$g(X,Y)$存在数学期望,则

$$Eg(X,Y) = \begin{cases} \sum_{i,j} g(x_i, y_j)p_{ij}, & 离散型，\\ \iint_{\mathbb{R}^2} g(x,y)p(x,y)dxdy, & 连续型。\end{cases}$$

- 若$g(x,y) = x$，则$EX = \begin{cases} \sum_{i,j} x_i p_{ij}, & 离散型，\\ \iint_{\mathbb{R}^2} xp(x,y)dxdy, & 连续型。\end{cases}$
- 特别地,若$g(x,y) = ax + by$，则$E(aX + bY) = aEX + bEY$.
- 一般地，$E\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i EX_i$。

12、二项分布的数学期望

例

求$X \sim b(n,p)$的数学期望。

解：

- 将X表示成如下的线性和

$$X = X_1 + \cdots + X_n, \qquad X_1, \ldots, X_n \overset{i.i.d}{\sim} b(1,p),$$

- 则$EX = EX_1 + \cdots + EX_n = np$。

注： 比如我们掷一个质量均匀分布的骰子，每掷一次，我们得到1点的可能性是$\frac{1}{6}$。如果我们独立地连掷100次，我们自然有理由期望得到$100 \cdot \frac{1}{6} \approx 17$次左右的1点。

13、匹配数的期望

例

将n个学生的学生证随机地分发给每个人，问平均有多少人恰好拿到自己的学生证？

解：

- 记X为恰好拿到自己学生证的人数。
- 记X_k为事件“第k个人恰好拿到自己的学生证”的示性函数.
- 则$X = X_1 + \cdots + X_n$且$X_k \sim b(1, \frac{1}{n})$（$\forall k$）。
- $EX = E(X_1 + \cdots + X_n) = EX_1 + \cdots + EX_n = n\frac{1}{n} = 1$。

注： 可以试一试用X的分布或$X_1, \ldots, X_n$的联合分布求EX。

问： X_k之间是独立的吗?

14、独立变量的特征数

性质

若X与Y相互独立，则

1. $E(XY) = EXEY$；
2. $Var(X \pm Y) = Var(X) + Var(Y)$。

证明：（1），只看离散型，连续型类似。

$$\begin{aligned} \mathrm{E}(XY) &= \sum_{i,j} x_i y_j P(X = x_i, Y = y_j) \\ &= \sum_{i,j} x_i y_j P(X = x_i) P(Y = y_j) \\ &= \sum_i x_i P(X = x_i) \sum_j y_j P(Y = y_j) \\ &= EXEY。\end{aligned}$$

15、独立变量和的方差

性质（2）

若X与Y相互独立，则$Var(X \pm Y) = Var(X) + Var(Y)$。

证明：

$$\begin{aligned} \mathrm{Var}(X \pm Y) &= E[X \pm Y - E(X \pm Y)]^2 \\ &= E[X - EX \pm (Y - EY)]^2 \\ &= E(X - EX)^2 \pm 2E[(X - EX)(Y - EY)] \\ &\quad + E(Y - EY)^2 \\ &= Var(X) \pm 2E(X - EX)E(Y - EY) + Var(Y) \\ &= Var(X) + Var(Y)。\end{aligned}$$

16、独立变量的特征数

推论

若$X_1, X_2, ..., X_n$相互独立，则

1. $E(X_1X_2\cdots X_n) = EX_1EX_2\cdots EX_n$；
2. $Var\left(\sum_{i=1}^n a_iX_i\right) = \sum_{i=1}^n a_i^2 Var(X_i)$。

例：求$X \sim b(n,p)$的方差。

- 将X表示成如下的线性和
  $X = X_1 + \cdots + X_n, \qquad X_1, \ldots, X_n \overset{i.i.d}{\sim} b(1,p)$，
- 则$Var(X) = \sum_{i=1}^n Var(X_i) = nVar(X_1) = np(1-p)$。
- 匹配数的方差呢？
- 答案和期望一样，居然也是常数1！你试试看？

17、协方差与相关系数

$Var(X+Y) = Var(X) + 2E[(X-EX)(Y-EY)] + Var(Y)$。

定义

- 对两个随机变量X, Y，称$E[(X-EX)(Y-EY)]$为X与Y的协方差（covariance）或相关（中心）矩，记为$\mathrm{Cov}(X,Y)$。
- 定义X与Y的（线性）相关系数$Corr(X,Y)$(或$r(X,Y)$)为：

$$r(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sigma(X)\sigma(Y)}.$$

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E[(X-EX)(Y-EY)] \\ &= E(XY - YEX - XEY + EXEY) \\ &= E(XY) - EXEY。\end{aligned}$$

18、Cauchy-Schwarz不等式

引理

若EX^2，$EY^2 < \infty$，则$[E(XY)]^2 \le EX^2EY^2$，且
"$=$" $\Leftrightarrow P(X=0) = 1$或$\exists a \in \mathbb{R}$，$\ni P(Y = aX) = 1$。

证明：

- $E(XY)$的存在性：
  $$|XY| \le X^2 + Y^2。$$
- $E(XY)$可看成为对角线上非负的对称双线性函数！

定义 内积$\langle X, Y\rangle := E(XY)$，（注：$EX^2 = 0 \Leftrightarrow P(X=0) = 1$）

- 内积空间的Cauchy-Schwarz不等式：
  $$|\langle X, Y\rangle| \le \|X\| \cdot \|Y\|。$$ ✌

19、最佳线性预报

回忆 $E(X-m)^2 = \min_a E(X-a)^2$，$m = ?$

- $E[(X-m)\cdot 1] = 0 \Rightarrow m = EX$。

考虑 $E[X - (\hat{a}Y + \hat{b})]^2 = \min_{a,b} E[X-(aY+b)]^2$，$(\hat{a}, \hat{b}) = ?$

- 在随机变量所构成的内积空间中，

$$\begin{aligned} E[X-(aY+b)] &= 0, \\ E\{Y[X-(aY+b)]\} &= 0. \end{aligned}$$

$\Rightarrow$ $EX - aEY - b = 0$，
$E(XY) - aEY^2 - bEY = 0$。

$\Rightarrow$ $\hat{a} = \mathrm{Cov}(X,Y)/Var(Y)$，$\hat{b} = EX - \hat{a}EY$ 且最佳逼近为

$$\hat{X} = \frac{\mathrm{Cov}(X,Y)}{Var(Y)}(Y - EY) + EX,$$

称这个一次函数为线性回归（*linear regression*）或回归直线（*regression line*）。

20、线性回归

- 这时最小误差为

$$\begin{aligned} E(X-\hat{X})^2 &= E\left[X-EX-\frac{\mathrm{Cov}(X,Y)}{Var(Y)}(Y-EY)\right]^2 \\ &= Var(X)-\mathrm{Cov}(X,Y)^2/Var(Y) \\ &= Var(X)[1-r(X,Y)^2], \end{aligned}$$

- 若$r(X,Y)>0$（<0，=0），则称X与Y正相关（*positively correlated*）（负相关*negatively correlated*，不相关*uncorrelated*）。
- 相关系数$r(X,Y)$反映了Y数值的变化对X的值的影响。
- 当$r(X,Y)=\pm1$时，最小误差为零，这时

$$\frac{X-EX}{\sqrt{Var(X)}}=r(X,Y)\frac{Y-EY}{\sqrt{Var(Y)}},$$

以概率1成立。此结果是统计学中线性回归分析理论基础。

21、协方差与相关系数

注：

在随机变量所构成的内积空间中，

- 协方差$\mathrm{Cov}(X,Y)$定义了向量$X-EX$和$Y-EY$的内积，
- 方差$Var(X)$是$X-EX$的长度的平方，
- 标准差$\sigma(X)$是$X-EX$的长度，
- 相关系数$r(X,Y)$确定了向量$X-EX$和$Y-EY$的夹角余弦，
- 期望EX是X在常数子空间上的投影。

定义

随机变量

$$X^*=\frac{X-EX}{\sqrt{Var(X)}}=\frac{X-EX}{\sigma(X)}$$

满足$EX^*=0$，$Var(X^*)=1$，称为X的标准化。

22、（协）方差（阵）和相关系数的性质

1. $\mathrm{Cov}(X,X)=Var(X)$；$|\mathrm{Cov}(X,Y)|\le\sigma(X)\sigma(Y)$，$|r(X,Y)|\le 1$.
2. 若$EX^2, EY^2<\infty$，则
$\mathrm{Cov}(X,Y)=E(XY)-EX\cdot EY$，$Var(X)=EX^2-(EX)^2$。
3. 若随机变量X与Y相互独立，则$\mathrm{Cov}(X,Y)=0$；若还知$Var(X), Var(Y)>0$，则$r(X,Y)$=0。
4. $\mathrm{Cov}(\cdot,\cdot)$是对称双线性函数，即$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$，
$\mathrm{Cov}(aX\pm bY,Z)=a\mathrm{Cov}(X,Z)\pm b\mathrm{Cov}(Y,Z)$。
5. $Var\left(\sum_{i=1}^n X_i\right)=\sum_{i,j=1}^n \mathrm{Cov}(X_i,X_j)=\sum_{i=1}^n Var(X_i)+2\sum_{1\le i<j\le n}\mathrm{Cov}(X_i,X_j)$.
6. 当$X_1,\ldots,X_n$两两不相关时，$Var\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n Var(X_i)$。
7. 随机变量$\mathbf{X}=(X_1,\ldots,X_n)$的协方差阵$\mathrm{Cov}(\mathbf{X})$是对称非负定矩阵，对角线元素分别是$X_1,\ldots,X_n$的方差。当$X_1,\ldots,X_n$独立时，协方差阵$\mathrm{Cov}(\mathbf{X})$是对角矩阵。

23、独立和相关

性质7： $\mathrm{Cov}\left(\sum_{i=1}^n a_iX_i, \sum_{i=1}^n b_iX_i\right)=(a_1,\ldots,a_n)\mathrm{Cov}(\mathbf{X})(b_1,\ldots,b_n)^T$.

性质3： 独立⇒不相关；但不相关$\not\Rightarrow$独立。

不相关$\not\Rightarrow$独立

Θ	0	$\pi/2$	π	$3\pi/2$
P	1/4	1/4	1/4	1/4

- $X=\cos\Theta, Y=\sin\Theta$，那么$EX=EY=0$，
- $\mathrm{Cov}(X,Y)=E(XY)=\frac{1}{2}E(\sin 2\Theta)=0$，即$r(X,Y)=0$，
- 但$X^2+Y^2=1$，不独立！

因为 $P(X=0)=P(Y=0)=1/2$；$P(X=Y=0)=0$。

24、超几何分布$h(n,N,M)$

例

设有N件产品，其中有M件次品。若从中不放回地随机抽取n件检验，记X_i为事件“第i次结果是次品”的示性函数，则抽检次品总数$X=\sum_{i=1}^n X_i$且$X\sim h(n,N,M)$。求EX和$Var(X)$。

- 注意本例为Pólya模型，$(X_{i_1},\ldots,X_{i_k})$和$(X_1,\ldots,X_k)$同分布。

$\Rightarrow$ $P(X_i=1)=P(X_1=1)=M/N\Rightarrow \forall i,\ X_i\sim b(1,M/N)$。

$\Rightarrow$ $EX_i=M/N$且$EX=\sum_i EX_i=nM/N$，
$Var(X_i)=\frac{M}{N}(1-M/N)=M(N-M)/N^2$。

- $P(X_i=X_j=1)=P(X_1=X_2=1)=M(M-1)/[N(N-1)]$，

$$\Rightarrow X_iX_j\sim\begin{pmatrix}0 & 1\\ 1-\frac{M(M-1)}{N(N-1)} & \frac{M(M-1)}{N(N-1)}\end{pmatrix}$$

$$\Rightarrow Cov(X_i,X_j)=E(X_iX_j)-EX_iEX_j=\frac{M(M-1)}{N(N-1)}-\frac{M^2}{N^2}=\frac{M(M-N)}{N^2(N-1)}$$

$$\Rightarrow Var(X)=n\frac{M(N-M)}{N^2}-2\binom{n}{2}\frac{M(M-N)}{N^2(N-1)}=n\frac{M(N-M)}{N^2}\frac{N-n}{N-1}。$$

25、标准正态分布

$(X,Y)\sim N(0,0,1,1,\rho)$：

$$\begin{aligned}
p(x,y) &= \frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}=\frac{1}{\sqrt{2\pi}}e^{-y^2/2}g(x,y),\\
g(x,y) &= \frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}};\quad (\sim N(\rho y,1-\rho^2))\\
Cov(X,Y) &= E(XY)-EXEY=\iint_{\mathbb{R}^2}xyp(x,y)dxdy\\
&= \int_{\mathbb{R}}y\frac{1}{\sqrt{2\pi}}e^{-y^2/2}\int_{\mathbb{R}}xg(x,y)dxdy\\
&= \rho\int_{\mathbb{R}}y^2\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy=\rho EY^2=\rho Var(Y)=\rho,\\
r(X,Y) &= \rho.
\end{aligned}$$

26、正态分布

性质：若$(X,Y)\sim N(0,0,1,1,\rho)$，则$X,Y$不相关与独立等价。

证明：若$\rho=0$，则

$$p(x,y)=\frac{e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}}{2\pi\sqrt{1-\rho^2}}=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\frac{1}{\sqrt{2\pi}}e^{-y^2/2}=p_X(x)p_Y(y).$$

推论

若$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则$X$与$Y$不相关与独立等价。

- 作标准化：令$X^*=(X-\mu_1)/\sigma_1$，$Y^*=(Y-\mu_2)/\sigma_2$，
- 则$(X^*,Y^*)\sim N(0,0,1,1,\rho)$。
- $r(X^*,Y^*)=Cov(X^*,Y^*)=\rho$。
- $r(X,Y)=Cov(\sigma_1X^*+\mu_1,\sigma_2Y^*+\mu_2)/(\sigma_1\sigma_2)=\rho$。
- X,Y不相关$\Leftrightarrow\rho=0\Leftrightarrow X^*,Y^*$相互独立$\Leftrightarrow X,Y$相互独立。