随机变量的函数

xyfJASON

1 一元函数：$Y = g(X)$

1.1 $\mathbf{PMF/PDF}$

设 X 是一离散随机变量，则 $Y = g(X)$ 也是一个离散随机变量，且其分布列为：

$$p_Y(y) = \sum_{\{x|y=g(x)\}} p_X(x)$$

设 X 是一连续随机变量，则 $Y = g(X)$ 也是一个连续随机变量，且其分布函数为：

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(g(X) \leqslant y) = \int\limits_{\{x|g(x)\leqslant y\}} f_X(x)\mathrm{d}x$$

于是 Y 的概率密度函数为：

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y}$$

1.2 期望

设 X 是一离散随机变量，则 $Y = g(X)$ 的期望为：

$$\mathbb{E}Y = \mathbb{E}[g(X)] = \sum_x g(x)p_X(x)$$

设 X 是一连续随机变量，则 $Y = g(X)$ 的期望为：

$$\mathbb{E}Y = \mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)\mathrm{d}x$$

也就是说，我们不必先求出 Y 的分布，只需知道 X 的分布就能求出 Y 的期望。

1.3 特殊情形

1.3.1 线性函数

设 X 是一连续随机变量，$a,b\in\mathbb{R}$ 且 $a\neq 0$，设 $Y=aX+b$，则：

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)$$

证：使用1.1节中描述的方法，先求 Y 的分布函数：

$$F_Y(y)=\mathbb{P}(Y\leqslant y)=\mathbb{P}(aX+b\leqslant y)=\begin{cases}\mathbb{P}\left(X\leqslant \frac{y-b}{a}\right)=F_X\left(\frac{y-b}{a}\right) & a>0\\ \mathbb{P}\left(X\geqslant \frac{y-b}{a}\right)=1-F_X\left(\frac{y-b}{a}\right) & a<0\end{cases}$$

然后求导得到 Y 的概率密度函数：

$$f_Y(y)=\frac{\mathrm{d}F_Y(y)}{\mathrm{d}y}=\begin{cases}\frac{1}{a}f_X\left(\frac{y-b}{a}\right) & a>0\\ -\frac{1}{a}f_X\left(\frac{y-b}{a}\right) & a<0\end{cases}=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)$$

证毕。

> 例子：正态分布的线性变换仍然是正态分布，且：
>
> $$X\sim N(\mu,\sigma^2)\implies aX+b\sim N(a\mu+b,a^2\sigma^2)$$
>
> 证：运用上述定理，
>
> $$\begin{aligned}f_{aX+b}(y)&=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right)\\&=\frac{1}{\sqrt{2\pi}\sigma|a|}\exp-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}\\&=\frac{1}{\sqrt{2\pi}\sigma|a|}\exp-\frac{(y-a\mu-b)^2}{2a^2\sigma^2}\end{aligned}$$
>
> 所以，$aX+b\sim N(a\sigma+b,a^2\sigma^2)$. 证毕。

1.3.2 单调函数

设 X 是一连续随机变量，g 是一严格单调的可逆函数，且其反函数 h 可微，则 $Y=g(X)$ 在 $\{y|f_Y(y)>0\}$ 内的概率密度函数是：

$$f_Y(y)=f_X(h(y))\left|\frac{\mathrm{d}h}{\mathrm{d}y}(y)\right|$$

证：仍然先求 Y 的分布函数：

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \begin{cases} \mathbb{P}(X \leqslant h(y)) = F_X(h(y)) & g \text{ 单调递增} \\ \mathbb{P}(X \geqslant h(y)) = 1 - F_X(h(y)) & g \text{ 单调递减} \end{cases}$$

于是求导得：

$$f_Y(y) = \begin{cases} f_X(h(y))h'(y) & g \text{ 单调递增} \\ -f_X(h(y))h'(y) & g \text{ 单调递减} \end{cases} = f_X(h(y) \left| \frac{\mathrm{d}h}{\mathrm{d}y}(y) \right|$$

证毕。

2 二元函数：$Z=g(X,Y)$

2.1 PMF/PDF

设 X,Y 是离散随机变量，则 $Z=g(X,Y)$ 也是一个离散随机变量，且其分布列为:

$$p_Z(z)=\sum_{\{(x,y)|z=g(x,y)\}} p_{X,Y}(x,y)$$

设 X,Y 是连续随机变量，则 $Z=g(X,Y)$ 也是一个连续随机变量，且其分布函数为:

$$F_Z(z)=\mathbb{P}(Z\leqslant z)=\mathbb{P}(g(X,Y)\leqslant z)=\iint\limits_{\{(x,y)|g(x,y)\leqslant z\}} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$

于是 Z 的概率密度函数为:

$$f_Z(z)=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}$$

2.2 期望

设 X 和 Y 是联合离散随机变量，则 $Z=g(X,Y)$ 的期望为:

$$\mathbb{E}Z=\mathbb{E}[g(X,Y)]=\sum_x\sum_y g(x,y)f_{X,Y}(x,y)$$

设 X 和 Y 是联合连续随机变量，则 $Z=g(X,Y)$ 的期望为:

$$\mathbb{E}Z=\mathbb{E}[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y$$

同样地，我们不必先求出 Z 的分布，只需知道 X 和 Y 的联合分布就能求出 Z 的期望。

2.3 特殊情形

2.3.1 独立随机变量和的分布

设 X,Y 是独立的离散随机变量，$Z=X+Y$，则：

$$p_Z(z)=\sum_x p_X(x)p_Y(z-x)$$

设 X,Y 是独立的连续随机变量，$Z=X+Y$，则：

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x$$

称上述两个式子为卷积（convolution）。

> 例子：相互独立的正态随机变量之和仍服从正态分布。
>
> 设 $X\sim N(\mu_x,\sigma_x^2)$，$Y\sim N(\mu_y,\sigma_y^2)$ 且相互独立，$Z=X+Y$，则 $Z\sim N(\mu_x+\mu_y,\sigma_x^2+\sigma_y^2)$.
>
> 证：运用上述定理，
>
> $$\begin{aligned} f_Z(z)&=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}-\frac{(z-x-\mu_y)^2}{2\sigma_y^2}}\mathrm{d}x\\ &=\int_{-\infty}^{+\infty}\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left[\frac{u^2}{\sigma_x^2}+\frac{(v-u)^2}{\sigma_y^2}\right]}\mathrm{d}u \qquad u=x-\mu_x,\, v=z-\mu_x-\mu_y \end{aligned}$$
>
> 由于
>
> $$\frac{u^2}{\sigma_x^2}+\frac{(v-u)^2}{\sigma_y^2}=\frac{u^2}{\sigma_x^2}+\frac{u^2}{\sigma_y^2}+\frac{v^2}{\sigma_y^2}-\frac{2uv}{\sigma_y^2}=\left(\frac{\sqrt{\sigma_x^2+\sigma_y^2}}{\sigma_x\sigma_y}u-\frac{\sigma_x v}{\sigma_y\sqrt{\sigma_x^2+\sigma_y^2}}\right)^2+\frac{v^2}{\sigma_x^2+\sigma_y^2}$$
>
> 令 $t=\frac{\sqrt{\sigma_x^2+\sigma_y^2}}{\sigma_x\sigma_y}u-\frac{\sigma_x v}{\sigma_y\sqrt{\sigma_x^2+\sigma_y^2}}$，则：
>
> $$f_Z(z)=\frac{1}{2\pi\sigma_x\sigma_y}\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}e^{\frac{v^2}{\sigma_x^2+\sigma_y^2}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}t^2}\mathrm{d}t=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}}e^{\frac{(z-\mu_x-\mu_y)^2}{\sigma_x^2+\sigma_y^2}}$$
>
> 所以 $Z\sim N(\mu_x+\mu_y,\sigma_x^2+\sigma_y^2)$. 证毕。
>
> 推论：设 $X_i\sim N(\mu_i,\sigma_i^2)$ 且相互独立，则 $\sum\limits_{i=1}^n X_i\sim N\left(\sum\limits_{i=1}^n\mu_i,\sum\limits_{i=1}^n\sigma_i^2\right)$.

2.3.2 极值分布

设 X, Y 是独立的随机变量，$M=\max(X, Y), N=\min(X, Y)$，则:

$F_M(z)=\mathbb{P}(M \leqslant z)=\mathbb{P}(X \leqslant z \wedge Y \leqslant z)=\mathbb{P}(X \leqslant z) \mathbb{P}(Y \leqslant z)=F_X(z) F_Y(z)$
$F_N(z)=\mathbb{P}(N \leqslant z)=1-\mathbb{P}(N>z)=1-\mathbb{P}(X>z \wedge Y>z)=1-\mathbb{P}(X>z) \mathbb{P}(Y>z)=1-[1-F_X(z)][1-F_Y(z)]$

更一般的，设 $X_1, X_2, \cdots, X_n$ 是 n 个相互独立的随机变量，$M=\max\limits_{1 \leqslant i \leqslant n} X_i$，$N=\min\limits_{1 \leqslant i \leqslant n} X_i$，则:

$$
\begin{aligned}
F_M(z) & =\prod_{i=1}^n F_{X_i}(z) \\
F_N(z) & =1-\prod_{i=1}^n\left[1-F_{X_i}(z)\right]
\end{aligned}
$$

2.3.3 瑞利分布

设 $X, Y \sim N(0, \sigma)$ 且相互独立，$R=\sqrt{X^2+Y^2}$，称 R 服从瑞利分布，其分布函数为:

$$
\begin{aligned}
F_R(r) & =\mathbb{P}(R \leqslant r)=\mathbb{P}\left(X^2+Y^2 \leqslant r^2\right) \\
& =\iint_{\left\{(x, y) \mid x^2+y^2 \leqslant r^2\right\}} p_{X, Y}(x, y) \mathrm{d} x \mathrm{d} y \\
& =\iint_{x^2+y^2 \leqslant r^2} \frac{1}{2 \pi \sigma^2} e^{-\frac{x^2+y^2}{2 \sigma^2}} \mathrm{d} x \mathrm{d} y \\
& =\frac{1}{2 \pi \sigma^2} \int_0^{2 \pi} \mathrm{d} \theta \int_0^r e^{-\frac{\rho^2}{2 \sigma^2}} \rho \mathrm{d} \rho \\
& =\left.e^{-\frac{\rho^2}{2 \sigma^2}}\right|_r^0 \\
& =1-e^{-\frac{r^2}{2 \sigma^2}}
\end{aligned}
$$

故概率密度函数为:

$$
f_R(r)=\frac{\mathrm{d} F_R(r)}{\mathrm{d} r}=\frac{r}{\sigma^2} e^{-\frac{r^2}{2 \sigma^2}} \quad(r>0)
$$

3 $\mathbf{Y} = T(\mathbf{X})$

设 $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ 是一个随机向量，T 是 $\mathbb{R}^n$ 上的一可逆映射，$\mathbf{Y} = T(\mathbf{X})$，则 $\mathbf{Y}$ 的概率密度函数为:

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(T^{-1}(\mathbf{y}))|J|$$

其中，J 表示 $T^{-1} : \mathbf{y} \mapsto \mathbf{x}$，即 $\mathbf{x} = T^{-1}(\mathbf{y})$ 的雅各比行列式:

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

证：设 $D \subset \mathbb{R}^n$ 是一个性质好的集合，则:

$$\begin{aligned} \mathbb{P}(\mathbf{Y} \in D) &= \mathbb{P}(T(\mathbf{X}) \in D) \\ &= \mathbb{P}(\mathbf{X} \in T^{-1}(D)) && \text{两边同时施以 } T^{-1} \\ &= \int_{T^{-1}(D)} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} \\ &= \int_D f_{\mathbf{X}}(T^{-1}(y))|J| \mathrm{d}\mathbf{y} && \text{变量代换 } \mathbf{y} = T(\mathbf{x}) \end{aligned}$$

又

$$\mathbb{P}(\mathbf{Y} \in D) = \int_D f_{\mathbf{Y}}(\mathbf{y}) \mathrm{d}\mathbf{y}$$

根据 D 一定的任意性，可知:

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(T^{-1}(\mathbf{y}))|J|$$

证毕。