随机变量及其分布

Didnelpsun

目录

1　二项分布　1

2　泊松分布　1

3　几何分布　1

4　均匀分布　2

5　指数分布　3

6　正态分布　4

1　二项分布

$P\{X=k\}=C_n^k p^k(1-p)^{n-k}$（$k=0,1,\cdots,n$，$0<p<1$），$X\sim B(n,p)$。

例题：已知随机变量 X 的概率密度为 $f(x)=\begin{cases}2x, & 0<x<1\\ 0,其他 & \end{cases}$，$Y$ 表示对 X 进行 3 次独立重复试验中出现事件 $\left\{X\leqslant\dfrac{1}{2}\right\}$，求 $P\{Y=2\}$。

解：已知对 X 进行独立重复试验，表示这个进行的是伯努利试验，从而 $Y\sim B(n,p)$。又是 3 次，所以 $Y\sim B(3,p)$。

只用求出这个 p 即 $\left\{X\leqslant\dfrac{1}{2}\right\}$ 的概率就可以了。又已知 $f(x)$。

$\therefore p=\left\{X\leqslant\dfrac{1}{2}\right\}=\int_0^{\frac{1}{2}}2x\,\mathrm{d}x=\dfrac{1}{4}$。$\therefore P\{Y=2\}=B\left(3,\dfrac{1}{4}\right)=\dfrac{9}{64}$。

2　泊松分布

$P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$（$k=0,1,\cdots,n$，$\lambda>0$），$X\sim P(\lambda)$。

例题：设一本书的各页印刷错误的个数 X 服从泊松分布。已知只有一个和只有两个印刷错误的页数相同，则随机抽查的 4 页中无印刷错误的概率 p 为？

解：$\because P\{X=1\}=P\{X=2\}$，$\therefore \dfrac{\lambda^1}{1!}e^{-\lambda}=\dfrac{\lambda^2}{2!}e^{-\lambda}$，$\lambda=2$。

由于随机抽四页类似于伯努利试验是相互独立的，所以随机抽 4 页都无错误的概率为 $[P\{X=0\}]^4=e^{-8}$。

3　几何分布

$P\{X=k\}=(1-p)^{k-1}p$（$k=0,1,\cdots,n$，$0<p<1$），$X\sim G(p)$。

例题：袋中有 8 个球，其中 3 个白球 5 个黑球，现在任意从中取出 4 个球，若四个球中有 2 个黑球和 2 个白球则试验停止，否则将其放回袋中重新抽取直到满足条件，用 X 表示试验次数，则求 $P\{X=k\}$。

解：由题目的停止，则说明这个题目的概率是服从几何分布的，最重要的就是求出单次满足事件概率 p。

根据组合和乘法原理，$p=\dfrac{C_3^2C_5^2}{C_8^4}=\dfrac{3}{7}$。

则 $P\{X=k\}=\left(\dfrac{4}{7}\right)^{k-1}\cdot\dfrac{3}{7}$。

例题：已知随机变量 X 的概率密度为 $f(x)=\begin{cases}2^{-x}\ln 2, & x>0\\ 0,\text{其他}\end{cases}$，对 X 进行独立重复观测，直到第 2 个大于 3 的观测值出现时停止，记 Y 为观测次数，求 Y 的概率分布。

解：由题目直到就停止，知道 $Y\sim G(p)$。

又 $p=P\{X\geqslant 3\}=\int_3^{+\infty}2^{-x}\ln 2\,\mathrm{d}x=\dfrac{1}{8}$

这是对几何分布的变形，首先进行 k 次试验，第 k 次成功，所以要乘 p，而因为是第 2 个成功，所以前面的 $k-1$ 次中有 $k-2$ 次失败和一次成功，所以一共 $p^2(1-k)^{k-2}$。因为前面的成功的一次在 $k-1$ 中任意一个地方就可以了，所以一共有 $k-1$ 中可能性，要考虑到排列，所以还要乘 $(k-1)$。

$\therefore P\{Y=k\}=(k-1)\left(\dfrac{1}{8}\right)^2\cdot\left(\dfrac{7}{8}\right)^{k-2}$。

4 均匀分布

$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases}$，$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leqslant x<b\\ 1, & x\geqslant b\end{cases}$，$X\sim U(a,b)$。

例题：已知随机变量 $X\sim U(a,b)$（$a>0$）且 $P\{0<X<3\}=\dfrac{1}{4}$，$P\{X>4\}=\dfrac{1}{2}$，求 X 的概率密度以及 $P\{1<X<5\}$。

解：$\because P\{X>4\}=\dfrac{1}{2}$，4 在其区间中点上，$\dfrac{a+b}{2}=4$。

$\because P\{0<X<3\}=\dfrac{1}{4}$，3 若在 a 左边则概率为 0，所以必然在右边。

$\therefore P\{a<X<3\}=\dfrac{1}{4}$，$P\{<3X<4\}=1-\dfrac{1}{4}-\dfrac{1}{2}=\dfrac{1}{4}$，$\dfrac{4-3}{b-a}=\dfrac{1}{4}$。

解得 $a=2$，$b=6$，$X\sim U(2,6)=f(x)=\begin{cases}\dfrac{1}{4}, & 2<x<6\\ 0, & \text{其他}\end{cases}$。

$P\{1<X<5\}=\dfrac{5-2}{6-2}=\dfrac{3}{4}$。

例题：已知随机变量 X 在区间 $[0,1]$ 上服从均匀分布，在 $X=x$（$0<x<1$）的条件下随机变量 Y 在区间 $[0,x]$ 上服从均匀分布。

(1)(X,Y) 的概率密度。

解：X 在区间 $[0,1]$ 上服从均匀分布，则 $X\sim f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}$。

Y 在 $X=x$ 下均匀分布，则 $f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{x}, & 0<y<x<1\\ 0, & 其他\end{cases}$。

(X,Y) 联合概率 = 条件概率 × 边缘概率。

即 $f(x,y)=f_{Y|X}(y|x)f_X(x)=\begin{cases}\dfrac{1}{x}, & 0<y<x<1\\ 0, & 其他\end{cases}$。

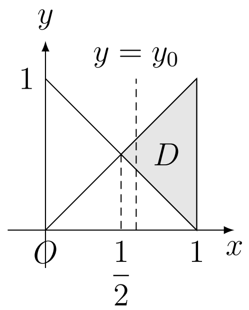
(2)Y 的概率密度。

解：首先求 Y 的边缘概率密度，就需要积 X。然后求 y 的区间，XY 的联合区间是横坐标 $[0,1]$ 到纵坐标 $[0,1]$ 的下三角形，则 $y\in[0,1]$。

然后求 Y 就在联合概率密度所规定的区间中画一条 $y=y_0$ 的线，从左先交到的是 $y=x$，所以下限就是 y，后交的是 $x=1$，所以上限为 1。最后将 y 的联合分布函数放在中间，得到 $f_Y(y)=\begin{cases}\displaystyle\int_y^1\dfrac{1}{x}\mathrm{d}x=-\ln y, & 0<y<x<1\\ 0, & 其他\end{cases}$。

(3) 概率 $P\{X+Y>1\}$。

解：求 $P\{X+Y>1\}$ 就是求一个区间的概率值，即 $P\{(X,Y)\in G\}=\displaystyle\iint_G f(x,y)\,\mathrm{d}x\mathrm{d}y$。



所以 $P\{X+Y>1\}=\displaystyle\iint_D\dfrac{1}{x}\mathrm{d}\sigma$，$D=x+y>1\cap 0<y<x<1$。$\displaystyle\iint_D\dfrac{1}{x}\,\mathrm{d}\sigma=\int_{\frac{1}{2}}^1\mathrm{d}x\int_{1-x}^x\dfrac{1}{x}\mathrm{d}y=1-\ln 2$。

5 指数分布

$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & 其他\end{cases}$，$F(x)=\begin{cases}1-e^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases}$，$X\sim E(\lambda)$。

例题：已知随机变量 $X\sim E(1)$，a 为常数且大于 0，求 $P\{X\leqslant a+1|X>a\}$。

解：$P\{X\leqslant a+1|X>a\}=\dfrac{P\{a<X\leqslant a+1\}}{P\{X>a\}}=\dfrac{\int_a^{a+1}e^{-x}\,\mathrm{d}x}{\int_a^{+\infty}e^{-x}\,\mathrm{d}x}=1-\dfrac{1}{e}$。

也可以根据指数分布的无记忆性：$P\{X\leqslant a+1|X>a\}=1-P\{X>a+1|X>a\}=1-P\{X>1\}=P\{X\leqslant 1\}=F(1)=1-\dfrac{1}{e}$。

例题：随机变量 X 服从参数为 1 的指数分布，求 $P\{3>X>2|X>1\}$。

已知 $F(X<x)=1-e^{-\lambda x}$，则 $F(X>x)=e^{-\lambda x}$。且 $2<X<3\cap 1<X=2<X<3$。

则 $P\{3>X>2|X>1\}=\dfrac{P\{3>X>2\}}{P\{X>1\}}=\dfrac{P\{X>2\}-P\{X>3\}}{P\{X>1\}}=\dfrac{e^{-2}-e^{-3}}{e^{-1}}=e^{-1}-e^{-2}$。

6 正态分布

$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$（$-\infty<x<+\infty$，$-\infty<\mu<+\infty$，$\sigma>0$），$X\sim N(\mu,\sigma^2)$。

例题：已知随机变量 $X\sim N(0,1)$，对给定的 α（$0<\alpha>1$），数 μ_α 满足 $P\{X>\mu_\alpha\}=\alpha$，若 $P\{|X|<x\}=\alpha$，求 x。

解：$P\{X>\mu_\alpha\}=\alpha$ 即表示 μ_α 为标准正态分布的上 α 分位点。

又 $P\{|X|<x\}=\alpha$，即 $-x<X<x$ 的面积为 α，所以两边的面积各为 $\dfrac{1-\alpha}{2}$，$P\{X<x\}=P\{X>x\}=\dfrac{1-\alpha}{2}$。

$\because$ 面积为 α 的下标为 α，$\therefore$ 面积为 $\dfrac{1-\alpha}{2}$ 的下标为 $\dfrac{1-\alpha}{2}$，$x=\mu_{\frac{1-\alpha}{2}}$。