连续随机变量

这章介绍连续随机变量下的分布。

均匀分布

均匀分布的问题是我前面讲的一个问题得来的，也就是在一段范围$[a,b]$内取一个值，那么这个值的概率是多少，很简单嘛，就是这段距离然后用一除。即$\frac{1}{b-a}$ 那么概率密度函数的值为

$$f_X(x)=\begin{cases}\frac{1}{b-a} & 若a\le x\le b\\ 0 & 其他\end{cases}$$

随机变量X服从区间$[a,b]$上的均匀分布，我们说X~$Unif(a.b)$
均值和方差分别是

$$\mu x=\frac{b+a}{2}，\ \sigma_X^2=\frac{(b-a)^2}{12}$$

均值的证明

$$\mu x=\int_{-\infty}^{\infty}x\frac{1}{b-a}dx=\frac{b+a}{2}$$

方差的证明

$$\sigma_X^2=\int_{-\infty}^{\infty}(x-\mu x)^2\frac{1}{b-a}dx=\frac{(b-a)^2}{12}$$

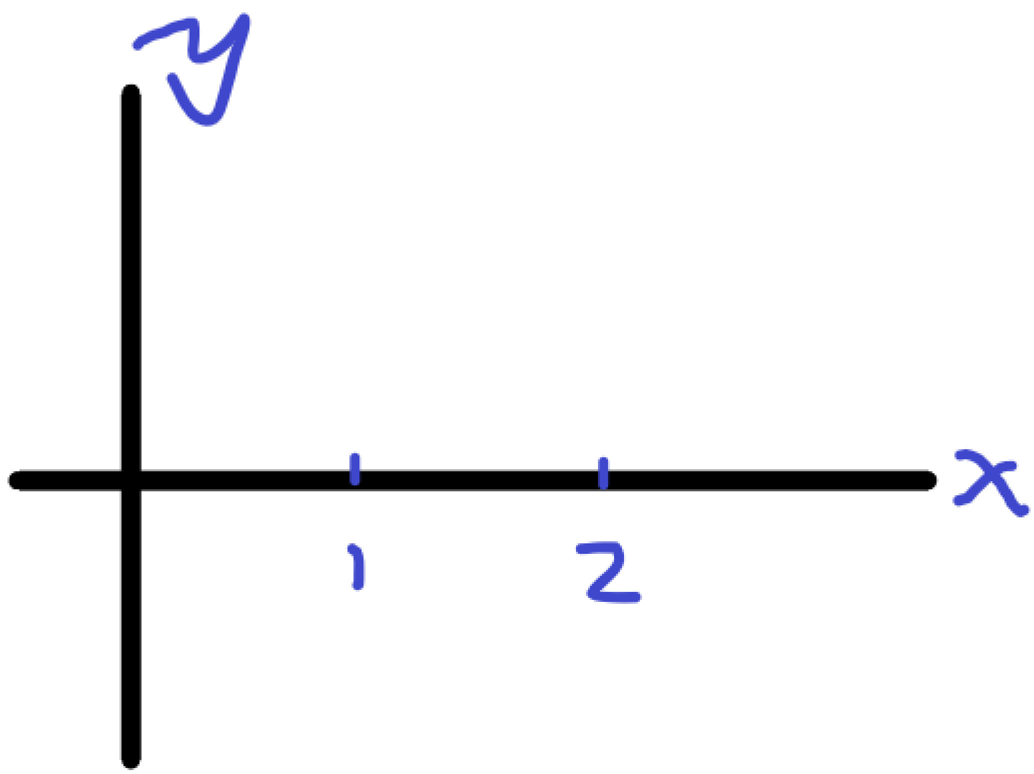
服从均匀分布的随机变量之和

我们应该想到前面第九章的卷积公式。输出一个特定值，那么变的就只有在变量里面取舍了。

设X，Y是相互独立的随机变量，并且都服从$Unif(0,1)$，那么$Z=X+Y$的概率密度函数是

$$f_Z(z)=\begin{cases} z & 若 0\le z\le 1 \\ 2-z & 若 1\le z\le 2 \\ 0 & 其他 \end{cases}$$

由于X，Y都在$[0,1]$上取值，考察极端情况，每个变量都有最大值为1的情况，那么应该有$max(X+Y)=2$，所以分成这三个区间是合理的。为什么是$2-z$呢，因为我们不考虑x,y小于1 的值，小于1的取值是无意义的。是一堆垃圾，不在我们的概率空间内。因为在我们分好的三个区域内，每个区域的概率是等同的，所以减1就使得大于1的值能够保持在一个一样的区域内。我们画个图吧。



所以我们把一个取值空间分段，**（忽略x，y这会误导你对问题的思考。我画坐标轴的习惯）**，然后我们在大于1上面的值取概率，那不就离谱了，超过1 的数字概率超过一，那就是不可能的了，所以用$2-1$表示这个区间。

我们就要开始用卷积的思想了，有先决条件X+Y = Z，所以有z的概率为$f_Z(z)=\int f(t)f(z-t)dt$，那我们看个例子吧

例子

我们设X，Y是相互独立且服从均匀分布的随机变量，那么$Z=X+Y$在$\frac{1}{3}$到$\frac{3}{2}$之间的取值是多少呢？，那么很简单，我们进行分类，把这个区间按照距离为1分段，那么有$\frac{1}{3}$到1和1到$\frac{3}{2}$到2
分别积分就行了

$$P(\frac{1}{2}\leq Z\leq\frac{3}{2})=\int_{1/2}^{3/2}f(z)dz$$

$$=\int_{1/2}^{1}f(z)dz+\int_{1}^{3/2}f(z)dz$$

$$=\int_{1/2}^{1}zdz+\int_{1}^{3/2}(2-z)dz=\frac{3}{4}$$

接下来我们还要复习前面的内容，例如概率的可加性，独立事件的概率，其实就是集合，求他们一起发生的可能性，就是并运算，也就是加法。那么我们可以把两个分段加起来就行了.

指数分布

如果随机变量X满足

$$f(x)=\begin{cases}\frac{1}{\lambda}e^{-x/\lambda} & 若x>0\\ 0 & 其他\end{cases}$$

我们讲，X是服从参数为$\lambda>0$的指数分布

那么他们的均值和方差分别是

$$\mu_X=\lambda,\ \sigma_X^2=\lambda^2$$

特别的，当$\lambda=1$，均值方差都是1。。

均值和方差的证明

这些东西都是通过求大量积分得到的.

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x) dx = \int^{\infty} x \frac{1}{\lambda} e^{-x/\lambda} dx$$

令 $t = x/\lambda$，那么积分变为

$$\mu_X = \lambda \int^{\infty} t e^{-t} dt$$

注意到这是一个伽马函数，如果你知道伽马函数，那么很简单就能算出来$\Gamma(2) = 1$

普遍一点的方法是利用分部公式

$$\mu_X = \lambda\big[-te^{t}|_0^{\infty} - \int^{\infty} e^{-t} dt\big] = \lambda$$

方差的证明会麻烦一点

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$$

$$\int_0^{\infty} \lambda^2 (\frac{x}{\lambda} - 1)^2 e^{-x/\lambda} \frac{dx}{\lambda}$$

$$use \quad t = x/\lambda$$

$$= \lambda^2 \int^{\infty} (t-1)^2 e^{-u} dt$$

$$use\, u \ = (t-1)^2 \ \ and \ \ dv = e^{-t} dt$$

$$So, we \quad have$$
$$du = 2(t-1)dt \ \ and \ \ v = -e^{-t}$$
$$because \quad use \quad the \quad int \quad parts \quad ruls, We \quad could$$

$$\sigma_X^2 = \lambda^2 [(t-1)^2 (-e^{-t}) \Big|_0^{\infty} + \int^{\infty} 2(t-1) e^{-t} dt] = \lambda^2$$

服从指数分布的随机变量之和

爱尔朗首先用这个来分析电话的分布，现在很广泛的用于排队论中。

这个理论很有用的阐述了对于n个相互独立且服从指数分布的随机变量，他们的和刚刚好就是这个分布。

设$X_1 \ldots X_n$是n个独立分布的随机变量，而且均服从参数为$\lambda > 0$ 的指数分布，那么$X = X_1 + X \ldots X_n$的概率密度函数是

$$f_X(x) = \begin{cases} \frac{x^{n-1} e^{-x/\lambda}}{\lambda^n (n-1)!} & x \geq 0 \\ 0 & \text{其他} \end{cases}$$

例子

指数分布一般用来表示独立事件随机发生的时间间隔，例如旅客进入机场的时间间隔、打进客服中心电话的时间间隔等等。

要明白的是，关于泊松分布，我们先来讲一下，从二次分布开始，我们假定在每个独立的时间（也就是我们设定的时间）发生了一件事情，4个时间段发生了三次事件，然后一次没发生。自然就是二项式分布

$$\binom{4}{3} p^3 (1-p)^1$$

但是如果一个区域内直接发生了多件事情，那么原有的时间段就不够分了，所以我们要继续分，分到满足二项式系数的定义。分到后面就是我们的泊松分布了。也就是说在T时间内发生了k件事情的概率就是

$$p(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

其中的λ就是我们的均值，$\lambda = np$，也就是发生了事件的平均次数

泊松分布可以描述一个标准时间段（我们自己规定）发生事件的次数的概率。但是在发生事件的之间的空时间段有多少呢？这个间隔又服从什么分布呢？

没有发生事件的概率由泊松分布得到

$$p(X = 0) = \frac{\lambda^0}{0!} e^{-\lambda} = e^{-\lambda}$$

拓展一下我们就可以得到泊松过程

$$P(X = k, t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

我们加入了一个新变量——t，用他来表明我们的时间段的数量，例如1就是我们的泊松分布，一个标准时间段，如果是$\frac{1}{2}$，那就是一半的时间段，t乘以λ就是要求的时间段的均值，这个东西我们也叫泊松过程，如果我们把$\frac{1}{t}$看成是一个标准时间段，那么就得出了我们书上定义的概率密度函数。

那么两次发生事情之间的间隔大于t的概率就是t时间内没有发生的概率，那就是

$$Y = \text{两次发生的事件之间的时间间隔}$$

$$P(Y > t) = P(X = 0, t) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}, t > 0$$

那么小于等于t时间的间隔就是，也就是标准化度量，用容斥定理得到

$$P(Y \geq t) = 1 - P(Y > t) = 1 - e^{-\lambda t}\text{，求一次导就是我们的概率密度函数}\lambda e^{-\lambda t}\text{，我们的教材是}\frac{1}{\lambda} e^{-x/\lambda}$$