Chpt.4 Digital Features of Random Variables

第四章 随机变量的数字特征

随机向量的函数的分布

$$Z = X + Y \qquad f_z(z) = \int_{-\infty}^{+\infty} f(x, z-x)dx = \int_{-\infty}^{+\infty} f(z-y, y)dy$$

$$Z = X - Y \qquad f_z(z) = \int_{-\infty}^{+\infty} f(x, x-z)dx = \int_{-\infty}^{+\infty} f(z+y, y)dy$$

$$f_{Y/X}(z) = \int_{-\infty}^{\infty} |x| f(x, xz)\mathrm{d}x,$$

$$f_{XY}(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f(x, \frac{z}{x})\mathrm{d}x.$$

$$F_z(z) = P\left(\max(X_1, X_2, \cdots, X_n) \le z\right) = F(z, z, \cdots, z)$$

$$F_z(z) = P\left(\min(X_1, \cdots, X_n) \le z\right) = 1 - P\left(X_1 \ge z, \cdots, X_n \ge z\right)$$

特殊分布

X服从参数α,β的伽玛分布，X~Γ(α, β)

- $\Gamma(1, \beta)$为指数分布
- $\Gamma\left(\frac{n}{2}, 2\right)$ 为参数为n的χ^2分布，记为$X \sim \chi^2(n)$

若 $X_1, \cdots, X_n$ 相互独立，都服从$N(0,1)$，则

$$Y = X_1^2 + \cdots + X_n^2 \sim \chi^2(n)$$

分布的参数可加性

$X \sim B(n_1, p), Y \sim B(n_2, p)$,两者独立，则 $X+Y \sim B(n_1+n_2, p)$

$X \sim \pi(\lambda_1), Y \sim \pi(\lambda_2)$,两者独立，则 $X+Y \sim \pi(\lambda_1+\lambda_2)$

若 X_1 X_2 相互独立， $X_1 \sim N(\mu_1, \sigma_1^2)$ $X_2 \sim N(\mu_2, \sigma_2^2)$

则 $X_1 + X_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

若X_1, X_2相互独立，$X_1 \sim \Gamma(\alpha_1, \beta)$, $X_2 \sim \Gamma(\alpha_2, \beta)$, 则

$X_1+X_2 \sim \Gamma(\alpha_1+\alpha_2, \beta)$

χ^2分布具有可加性，设$X_1 \sim \chi^2(n_1)$, $X_2 \sim \chi^2(n_2)$，

则$X_1+X_2 \sim \chi^2(n_1+n_2)$

Why Digital Features?

Reviews

引入随机变量$X: \Omega \to \mathbb{R}$，把概率研究抽象化、数字化

引入分布函数$F: \mathbb{R} \to \mathbb{R}$，从而可以借鉴已有的数学工具

WHY Digital Features

[1]分布函数全面反映随机变量的性质，但是并不易掌握。

随机变量是否有一些特征，比较突出地反映其性质？

如取所有值的平均情况，X与均值的偏离情况。

例：研究生入学考试成绩

- 限定总分，由总分可以知道多科的平均分；
- 但是同时限定单科的最低分；
- 平均值高、偏离均值小的才是好的。

Why Digital Features?

[2] 当我们比较不同的随机变量时，分布函数或分布律并不可用，以分布律而言，它只说明变量取某值的可能性，是概率特性，从概率的角度我们不能说一个变量比另一个变量如何。

Example (pp.81, 例5) 设系统L由两个相互独立的子系统L_1, L_2链接而成，连接的方式分别为（1）串联（2）并联（3）备用

串联min(X,Y), $f_1(z)=(\alpha+\beta)e^{-(\alpha+\beta)z}$

并联Max(X,Y)， $f_2(z)=\alpha e^{-\alpha z}+\beta e^{-\beta z}-(\alpha+\beta)e^{-(\alpha+\beta)z}$

备用X+Y， $f_3(z)=\dfrac{\alpha\beta}{\beta-\alpha}[e^{-\alpha z}-e^{-\beta z}]$

直观上看，应该是系统寿命：备用 ＞ 并联 ＞ 串联；

但是，由于是随机的，可能具体的一个备用模式的寿命比其他还短

Why Digital Features?

[3] 根据经验，某些随机现象的随机变量服从某类分布，它们的一些参数可由某些数字特征确定. 对这些随机现象，数字特征有更重要的意义.

E.g. 正态分布 $N(\mu, \sigma^2)$，二项分布 $B(n, p)$，泊松分布$\pi(\lambda)$

主要的数字特征

- 数学期望：描述一个随机变量的平均水平（即均值）
- 方　　差：描述一个随机变量的相对于均值的离散程度
- 相关系数：描述两个随机变量间线性关系的密切程度

分赌本问题

甲、乙二人各出同等数目的赌注，比方说1000元，然后进行博弈。每一局甲胜和乙胜都有同等概率，即0.5。二人约定：谁先胜满m局(如在7盘棋中，m=4），谁就取走全部赌注2000元。

到某时为止，甲胜了 a 局，而乙胜 b 局，a、b都比 m 小，而二人因故要中止博弈，问这2000元赌注该如何分才算公平。

分析：

[1] 如果a＝b，则二人态势一样，平分赌注不会有异议；

[2] 若a、b不等，例如 a＞b，则甲理应多分一些才公平。

但具体的比例该如何才算公平?

4.1 随机变量的数学期望

若干不同的解法

[1] 帕西奥利（ 1494年）

按a：b的比例分配

[2] 塔泰格利亚（1556年）

按m＋a－b：m－a＋b的比例分配。

同时，他又指出，此问题最好交由法官判决。

[3] 法雷斯泰尼（1603年）

按2m－1＋a－b：2m－1－a＋b的比例分配

[4] 卡丹诺（1539年）

记u＝m－a，v＝m－b，其中u、v是甲、乙最终取胜尚需的胜局数，若a＞b．则u＜v。

按v(v＋1)：u(u＋1)的比例分配赌注。

[5] 巴斯卡（1654年）

假这两个人继续赌下去，

则至多不超过2m-1-(a+b)局，就会最后见分晓。

甲、乙最后获胜都有一定的概率，分别记为p和q (p＋q＝1)。

巴斯卡主张赌注应接p：q的比例去分配。

一般情况下：

比赛2n＋1盘，先胜n＋1盘者获胜；进行了a＋b局后终止比赛时，只要没有哪位选手胜了n＋1盘，后面理论上还要进行（2n＋1）-（ a＋b ）盘比赛。

[5] 巴斯卡（1654年）

假设a＝2，b＝1，

例如：假设7局4胜，则理论上后4局还需要赌。

甲获胜的可能为:（$C_4^2+C_4^3+C_4^4$）$*(1/2)^4$

乙获胜的可能为:（$C_4^3+C_4^4$）$*(1/2)^4$

故甲、乙最终获胜的概率分别为11：5。

赌本应按11：5的比例分给甲和乙。

例如：假设9局5胜，则理论上后6局还需要赌。

甲获胜的可能为:（$C_6^3+C_6^4+C_6^5+C_6^6$）$*(1/2)^6$

乙获胜的可能为:（$C_6^4+C_6^5+C_6^6$）$*(1/2)^6$

故甲、乙最终获胜的概率分别为42：22。

赌本应按21：11的比例分给甲和乙。

分赌本问题若干不同的解法

[5] 巴斯卡（1654年）

为什么按最终获胜的概率之比去分配是公平的？

二人按各自的期望所得去分配应是合理的。各人的“期望所得”：

在最后4局未赌之前，甲有可能获得2000元，但机会只有11 / 1 6，因而他的“期望所得”只有2000 * (11/16)元＝1375元。

同样，乙的“期望所得”为2000 * (5/1 6)元＝625元。

期望所得之比，与获胜概率之比是一回事，所以按最终获胜的概率之比去分配是公平的。

由此，引出了期望(Expectation)概念

4.1 随机变量的数学期望

Example 为评价甲的射击技术，随机观察甲的10次射击，统计各次击中环数x_k和频数v_k如下表，求他每次射击击中的平均环数. (其中$N=\Sigma v_k=10$).

击中环数x_k	8	9	10
频数v_k	2	5	3
频　率$f_k=v_k/N$	0.2	0.5	0.3

解 这是一般求平均数 $\overline{x}$ 问题，此时

$$\overline{x}=\left(8\times 2+9\times 5+10\times 3\right)/10$$
$$=8\times\frac{2}{10}+9\times\frac{5}{10}+10\times\frac{3}{10}$$
$$=9$$

写成一般形式就是

$$\bar{x} = \left(\sum_k x_k v_k\right) \Big/ N = \sum_k x_k\, v_k / N = \sum_k x_k f_k$$

若要全面考察甲的射击技术，仅凭这10次观察是不够的，因为频率 $v_k / N = f_k$ 是与该次射击(试验)的结果有关的.

观察次数N不断增多，由于频率稳定于概率 p_k ，和式 $\sum\limits_k x_k f_k$ 就趋于稳定 $\sum\limits_k x_k f_k \longrightarrow \sum\limits_k x_k p_k$

它是一个确定值，不依赖于具体试验，更能表示甲的射击水平.

4.1 随机变量的数学期望

4.1.1 数学期望定义

离散随机变量的均值

[Definition] 设离散型随机变量X的分布列为

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_k & \cdots \\ p_1 & p_2 & \cdots & p_k & \cdots \end{pmatrix}$$

如果级数$\sum x_k p_k$绝对收敛，即$\sum_k |x_k| p_k < \infty$，就称$\sum x_k p_k$为X的数学期望(mathematical expectation)或均值(mean)，记作E(X)。

Example 退化分布P{X=a}=1的数学期望 EX=a, 也即常数的数学期望就是它本身.

Example 二项分布X~B(n,p)的数学期望

$$E(X)=\sum_{k=0}^{n}kC_n^k p^k(1-p)^{n-k}=\sum_{k=1}^{n}kC_n^k p^k(1-p)^{n-k}$$

$$=\sum_{k=1}^{n}k\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

$$=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}(1-p)^{(n-1)-(k-1)}$$

$$=np\sum_{r=0}^{n-1}\frac{(n-1)!}{r!(n-1-r)!}p^{r}(1-p)^{(n-1)-r}$$

$$=np(p+[1-p])^{n-1}$$

$$=np$$

Remark 1： 均值的含义由上述例子可见。n 次试验中出现的平均数是np。

Remark 2：当n＝1时，两点分布 (0—1分布)的数学期望为p.

Remark3：条件 $\sum_k |x_k| p_k < \infty$ 表示 $E(X) = \sum x_k p_k$ 不受求和次序的影响，从而E(X)是一个确定的值（这是需要的，因实际情况中出现可能无次序的）. 当条件不满足时，就称X的数学期望不存在.

例 设随机变量X取值为 $(-1)^k 2^k / k$ (k = 1, 2, ….),

取得各值的概率为 $P\{X = (-1)^k 2^k / k\} = \dfrac{1}{2^k}$

由 $\sum_{k=1}^{\infty} |x_k| p_k = \sum_{k=1}^{\infty} \dfrac{1}{k} = \infty$(调和级数), $\sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} (-1)^k \dfrac{1}{k} < \infty$

(莱布尼兹级数)，仍说E(X) 不存在。

Example 泊松分布 $P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}\,(k=0,1,2,\cdots)$ 的数学期望

$$E(X)=\sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\sum_{k=1}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}$$

$$=\lambda\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}$$

概率之和为1

$$=\lambda*1$$

$$=\lambda$$

故泊松分布中参数λ就是它的数学期望.

Example 几何分布 $P\{X=k\}=p(1-p)^{k-1}(k=1,2,\cdots)$ 的数学期望。

$$E(X)=\sum_{k=1}^{\infty}kp(1-p)^{k-1}$$

记 q＝1－p

$$E(x)=p\frac{d}{dq}\left[\sum_{k=1}^{\infty}q^k\right]$$

$$=p\frac{d}{dq}\left[\frac{q}{1-q}\right]$$

$$=p\frac{(1-q)+q}{(1-q)^2}$$

$$=\frac{1}{p}$$

Example （关于概率、期望与策略选择） 圣彼得堡赌博问题：

1738年，Daniel Bernoulli 提出，论文有圣彼得堡科学院发表：

对于一个明智的赌徒，如果庄家需要他投注一定的金额，才能参加下列游戏，他的心理价位应该是多少？

1）投掷均匀硬币（p=0.5），head则中奖，本次游戏结束，tail则继续，直到出现head。

2）假设第k次中奖，则其奖金为 2^k

解 显然越晚中奖越好，我们尝试计算这个游戏的回报期望：

$\boldsymbol{k}$	1	2	3	…	k	…
$x=k$概率	½	¼	1/8	…	$0.5^{k-1}\cdot 0.5\ =0.5^k$	…
回报	2	4	8	…	2^k	…

Example （关于概率、期望与策略选择） 圣彼得堡赌博问题：

所以最终期望为

$$E(X)=\sum_{k=1}^{\infty} 2^k \times \left(\frac{1}{2}\right)^k=\sum_{k=1}^{\infty} 1=\infty$$

所以，也许你会支付4美元来玩这个游戏。当然，你的回报偶尔会非常大——8美元、16美元、32美元、64美元……理论上，回报可能接近无穷大。但是你会支付多少？这就是64美元的问题。

4.15.2 统计分析

一些社会科学研究人员认为，大多数人会花4美元来玩这个游戏，可能还会多一点。很少有人会出太多钱玩这个游戏。但是，从统计学角度分析，结果会是怎样的呢？你最多应该出多少钱？

好吧，我考虑上交我的统计粉丝俱乐部会员卡，因为我怕告诉你正确答案。由于涉及赌博，概率的规则建议人们应该不惜一切代价玩这个游戏。是的，一个统计学家会告诉你应不惜一切去玩这个游戏！只要成本没有达到无穷大，从理论上说，这就是一个好的赌注。

让我们算算。下表是前6次硬币翻转的回报：

翻转	可能性	游戏比例	赢得	预期支付
1	1：2	0.50	2美元	1美元
2	1：4	0.25	4美元	1美元
3	1：8	0.125	8美元	1美元

Example （关于概率、期望与策略选择*） 圣彼得堡赌博问题：

如何理解这个无穷大的期望：

不断多的高的入场费，都有机会通过在一个很大的k次情况下挣回来

这个结论是违背常识的

—理解期望

实际上，上述关于期望的结论说明：

对于一个非常高的入场费，*需要玩足够多次，才可以碰到一次挣大钱的机会，这是由于存在一个无穷大的回报所致。*

也就是说，哪怕成本是1,000,000,000,000

也可以通过进行 1000000+次 博一次 ∞的机会

成本、损失与策略：

在上述描述中，概率模型是不变的，这个概率模型的数字特征（期望）也没变，但是它不足以描述我们的目标，此时研究的是如何在这个概率下执行某种策略（规划、优化、统计等不同研究方法）。统计学策略：用于足够本金的玩家（庄）天生处于优势&赌博局的金额封顶策略。

Example （pp.93例5） 在一个人数很多的单位中普查某疾病, N个人验血，可用两种方法:

(1) 每个人化验一次，共需化验N次；

(2) k个人的血混合化验，如果结果是不含该病菌，说明这k个人都无该病，这样k个人化验一次即可；如果结果是含该病菌，则该组每个人再分别化验一次， k个人共化验 (k+1) 次.

试说明当生病概率p较小时，选取适当的k，按第二种方法可以减少化验次数，并说明k取什么值时最适宜。

解 记用第二种方法化验时，分N人为(N/k) 组。一组k人所需化验次数为X.

此时，如果都没有病，则k个人只化验一次，X=1概率为

$$P\{X=1\}=(1-p)^k$$

4.1 随机变量的数学期望

或者k个人至少有一人有病，需化验 (k+1) 次, 其概率为

$$P\{X=k+1\}=1-(1-p)^k$$

X	1	k+1
p	$(1-p)^k$	$1-(1-p)^k$

所以 $$E(X)=(1-p)^k+(k+1)[1-(1-p)^k]$$

$$=k+1-k(1-p)^k$$ 一个组测试数量的期望

一共有（N/k）组所需的平均化验次数为

$$\frac{N}{k}E(X)=N\left[1-(1-p)^k+\frac{1}{k}\right]$$

由此只要选择k使得下式成立，分组就比不分组节省次数。

$$N\left[1-(1-p)^{k}+\frac{1}{k}\right]< N$$

使得 $L=1-(1-p)^{k}+\frac{1}{k}<1$, 要 $k(1-p)^{k}>1$ 。

注意到p较小时，1－p 较大，容易取到 k 使 $k(1-p)^{k}>1$。

如果 p=0.1

问：p=0.01,0.001时会怎么样？

k	1	2	3	4	5	6	7
$k(1-p)^{k}$	0.9	1.62	2.187	2.6244	2.95425	3.1886	3.3480
$1-(1-p)^{k}+1/k$	1.1	0.69	0.6043	0.5939	0.60951	0.6352	

可见取 k=4 最好。如果N=1000，第二种方法平均化验数为

$$N\left[1-(1-p)^{k}+\frac{1}{k}\right]\approx 594$$

连续型随机变量的数学期望

连续型随机变量X，其密度函数为*f*(x). 其数学期望的定义可借鉴离散型情形的定义及普通积分的导出过程来引入.

STEP 1：先设X只在有限区间 [a, b] 上取值，将 [a, b]作分割：

$$a = x_0 < x_1 < \cdots < x_n = b$$

X落在各小段的概率

$$P\{x_k < X \le x_{k+1}\} = \int_{x_k}^{x_{k+1}} f(x)dx \approx f(x_k)\Delta x_k$$

近似地视为落在一点上x_k 的概率

STEP 2：与它相应的离散型随机变量的数学期望 $\sum x_k f(x_k)\Delta x_k$

STEP 3：令n→∞， $\max\{\Delta x_k\} \to 0$, 则和式的极限为积分.

$$\int_a^b xf(x)dx$$

连续型随机变量的数学期望

自然把它作为X的数学期望. 如果X在整个实轴(-∞, ∞)上取值，让a→∞，b→∞, 就得如下的一般定义.

Definition 设X为连续型随机变量，有密度函数f(x)，则当

$$\int_{-\infty}^{\infty} |x| f(x)dx < \infty$$

时，称 $E(x) = \int_{-\infty}^{\infty} xf(x)dx$ 为X的数学期望.

如果 $\int_{-\infty}^{\infty} |x| f(x)dx = \infty$，就称X的数学期望不存在.

Example 均匀分布的密度函数：$a \le x \le b$ 时，它的数学期望

$$E(x) = \int_a^b \frac{x}{b-a}dx = \frac{a+b}{2}$$

Example 正态分布 $X \sim N(\mu,\sigma^2)$ 的数学期望.

$$E(X)=\int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\sigma\int_{-\infty}^{\infty}\left\{\left(\frac{x-\mu}{\sigma}\right)+\frac{\mu}{\sigma}\right\}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}d\left(\frac{x-\mu}{\sigma}\right)$$

$$=\sigma\int_{-\infty}^{\infty}[y+\frac{\mu}{\sigma}]\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$$

$$=\mu\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy+\sigma\int_{-\infty}^{\infty}\boxed{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$$

$$=\mu$$

因此正态分布$N(\mu,\sigma^2)$ 中参数 μ 表示它的均值，这在密度函数的图形中已显示出来了.

Example 指数分布的数学期望

设随机变量X服从指数分布，其概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$

其中λ>0是常数. 求：E(X)

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx$$

$$=\int_0^{\infty}x\lambda e^{-\lambda x}dx$$

$$=\int_0^{\infty}-xde^{-\lambda x}$$

$$=-xe^{-\lambda x}\Big|_0^{\infty}+\int_0^{\infty}e^{-\lambda x}dx$$

$$=\int_0^{\infty}e^{-\lambda x}dx$$

$$=-\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{\infty}$$

$$=\frac{1}{\lambda}$$

如果密度函数表示为 $f(x)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & x>0\\ 0 & x\le 0\end{cases}$ $\quad E(x)=\theta$

Example 顾客平均等待多长时间?

设顾客在某银行的窗口等待的服务的时间 X(以分计)服从指数分布,其概率密度为

$$f(x)=\begin{cases}\frac{1}{5}e^{-x/5}, & x>0,\\ 0, & x\le 0.\end{cases}$$

试求顾客等待服务的平均时间?

解

$$E(X)=\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x=\int_{0}^{\infty} x\cdot\frac{1}{5}e^{-x/5}\,\mathrm{d}x=5(\text{分钟}).$$

因此,顾客平均等待5分钟就可得到服务.

Example 设随机变量X服从Γ(α, β)分布，求它的数学期望

解
$$f(x)=\begin{cases}0, & x\le 0\\ \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta} & x>0\end{cases}$$

$$\Gamma(a)=\int_0^{+\infty}u^{a-1}e^{-u}du$$

$$EX=\int_0^{+\infty}\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha}e^{-x/\beta}dx$$

$$\boxed{y=x/\beta}$$

$$=\int_0^{+\infty}\frac{\beta\, y^{\alpha}}{\Gamma(\alpha)}e^{-y}dy$$

$$=\frac{\beta\,\Gamma(\alpha+1)}{\Gamma(\alpha)}$$

$$=\alpha\beta$$

$$\Gamma(a+1)=\int_0^{+\infty}u^{a}e^{-u}du$$

$$=-\int_0^{+\infty}u^{a}de^{-u}$$

$$=-u^{a}e^{-u}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-u}du^{a}$$

$$=a\int_0^{+\infty}u^{a-1}e^{-u}du$$

$$=a\Gamma(a)$$

4.1.2 随机变量函数的数学期望

我们知道X的分布列为

X	$x_1 \cdots\cdots\cdots x_k$
p	$p_1 \cdots\cdots\cdots p_k$

可以知道 X^2 的分布列为

X^2	$x_1^2 \cdots\cdots\cdots x_k^2$
p	$p_1 \cdots\cdots\cdots p_k$

故此 $E(X^2)=\sum x_k^2 p_k$

函数g(x)，随机变量Y=g(x)，有 $E(g(X))=\sum g(x_k)p_k$

对于连续随机变量X,其概率密度函数为f(x)，则

$$\boxed{E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx}$$

上式在计算某些随机变量的数学期望时很有用. 如果Y=g(X)，我们不必求Y的概率密度函数 $f_y(y)$，再计算 $E(Y)=\int_{-\infty}^{\infty} yf_y(y)dy$ ，而是直接得到 $E(Y)=E(g(X))=\int_{-\infty}^{\infty} g(x)f(x)dx$

设有联合概率密度函数 $f(x_1,\cdots,x_n)$，$Y=g(X_1,\cdots,X_n)$，则

$$E(Y)=\iint\limits_{R^n} g(x_1,\cdots,x_n)f(x_1,\cdots,x_n)dx_1\cdots dx_n$$

特别地有

$$E(X_i)=\iint\limits_{R^n} x_i f(x_1,\cdots x_n)dx_1\cdots dx_n$$

Example 假X,Y是在区域 $\Omega=\{(x,y)\,|\,x^2+y^2\le 1\}$ 上满足均匀分布

求 E(X).

$$f(x,y)=\begin{cases}\dfrac{1}{\pi} & (x,y)\in\Omega \\ 0 & others\end{cases}$$

快速计算结果

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy=\iint\limits_{x^2+y^2\le 1} x\frac{1}{\pi}dxdy=0$$

Example 设$X\sim N(0,1), Y\sim N(0,1)$, X 与Y相互独立, 求 $E(\sqrt{X^2+Y^2})$

解: 我们没有必要去求 $g(X,Y)=\sqrt{X^2+Y^2}$ 的分布函数，只要求得X,Y的联合分布函数 $f(x,y)$即可。

注意到X,Y分别是标准正态分布，又是独立的，则联合分布为N(0,0,1,1)，对应的分布密度为$f(x,y)$

$$f(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi}\exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

$$E(\sqrt{X^2+Y^2})$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}dxdy$$

$$=\int_0^{2\pi}d\theta\int_0^{+\infty}\frac{1}{2\pi}\cdot r^2e^{-\frac{r^2}{2}}dr$$

$$=\int_0^{+\infty}r^2e^{-\frac{1}{2}r^2}dr$$

$$=\frac{\sqrt{2\pi}}{2}$$

Example （pp.81,例5）

串联 min(X,Y), $f_1(z)=(\alpha+\beta)e^{-(\alpha+\beta)z}$

并联 Max(X,Y)，$f_2(z)=\alpha e^{-\alpha z}+\beta e^{-\beta z}-(\alpha+\beta)e^{-(\alpha+\beta)z}$

备用 X+Y， $f_3(z)=\dfrac{\alpha\beta}{\beta-\alpha}[e^{-\alpha z}-e^{-\beta z}]$

比较三者的分布函数或概率密度没有意义，比如尽管说备用比串联好，但是这里是随机现象，而不是具体的两个元器件，很可能实际结果是第三种寿命比第一种还短。要从均值角度描述比较。

$$E(Min(X,Y))=\int_0^{\infty}z(\alpha+\beta)e^{-(\alpha+\beta)z}dz$$

$$=-ze^{-(\alpha+\beta)z}\Big|_0^{\infty}+\int_0^{\infty}e^{-(\alpha+\beta)z}dz$$

$$=-\frac{1}{\alpha+\beta}e^{-(\alpha+\beta)z}\Big|_0^{\infty}=\frac{1}{\alpha+\beta}$$

Example （pp.81,例5）

$$E(Max(X,Y)) = \int_0^\infty z\{\alpha e^{-\alpha z} + \beta e^{-\beta z} - (\alpha+\beta)e^{-(\alpha+\beta)z}\}dz$$

$$= \frac{1}{\alpha} + \frac{1}{\beta} - \frac{1}{\alpha+\beta}$$

$$E(X+Y) = \int_0^\infty z\frac{\alpha\beta}{\beta-\alpha}[e^{-\alpha z} - e^{-\beta z}]dz$$

$$= \frac{\beta}{\beta-\alpha}\frac{1}{\alpha} - \frac{\alpha}{\beta-\alpha}\frac{1}{\beta}$$

$$= \frac{1}{\alpha} + \frac{1}{\beta}$$

$M_3 > M_2 > M_1$

思考题： 国际市场上每年对我国某种商品的需求量X服从 [2000, 4000]上的均匀分布.每售出一吨该商品可获利3万美元; 但若积压于仓库，每吨将损失1万美元. 问应组织多少货源才能使收益最大?