# 组合分析

## 「计数基本原则」

- m种结果可造成事件1，对每种能造成事件1的结果，又有n种结果造成事件2，则对这两个事件共有mn种可能结果

## 「排列」

- 排列数：n!
- 有序则乘它，消序则除以它

## 「组合」

- 组合数：$\binom{n}{r}=\frac{n!}{(n-r)!r!}$
- pascal恒等式：$\binom{n}{r}=\binom{n-1}{r-1}+\binom{n-1}{r}$，分析为
  - 有$\binom{n-1}{r-1}$个子集含某特定元素，有$\binom{n-1}{r}$个子集不含某特定元素
- 二项式定理：$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^ky^{n-k}$

### 例题

#### 不同时入选

5女7男组成2女3男的委员会，其中有两个男的不同时入选，则共有多少种可能？
**用基本原则：女可能数*男可能数**
女可能数为：$\binom{5}{2}=10$，男可能数为：$\binom{7}{3}-1\times\binom{5}{1}=30$.则共300种。

#### 不相邻：插空法

n个东西，m个失效了，其他的没失效。n个东西都不能相互区分。问有多少种线性排列方式可以让任意两个失效的不相邻？

插空法：^1^1^1...^1^1^，1代表没失效的，^代表空，最多放一个失效的进去。
有插空方法：$\binom{n-m+1}{m}$

## 「多项式系数」

- 将n个物品分成r个可区分的组，第i组的容量为$n_i$，则共有$\frac{n!}{n_1!n_2!\cdots n_r!}=\binom{n}{n_1n_2\cdots n_r}$种分法
- 如果这r个组并无区别、不可区分，则应该在上面的基础上再除以$r!$，对这r个组进行消序
  - 如将10人分成2个5人队打篮球，则为$\binom{10}{5\ 5}/2!$

## 「方程整数解个数」

解方程$x_1+x_2+\cdots+x_r=n$，即为n个0要分成r组
用隔板法：0^0^0^0^...^0^0，共n-1块隔板，选r-1块则可，故共$\binom{n-1}{r-1}$个正整数解

# 概率论公理

## 「样本空间和事件」

- 样本空间：一次试验的所有可能结果构成样本空间
- 样本空间中的一个子集是一个事件
- 交：两事件同时发生
  - 若两事件的交是空集，称两事件是互斥的(mullually exclusive)
- 并：两事件至少一个发生
- 德摩根律：$\cup_1^nE_i^c=(\cap_1^nE_i)^c$

## 「概率论公理」

定义S是样本空间，E是事件。

$$Axiom\ 1:0\le P(E)\le 1$$
$$Axiom\ 2:P(S)=1$$
$$Axiom\ 3:若E_1,E_2,\cdots是互斥的，则有P(\cup_{i=1}^{\infty}E_i)=\sum_{i=1}^{\infty}P(E_i)$$

- 有人用频率的极限来定义概率，问题是频率极限一定收敛吗？若将频率极限一定收敛作为公理显然太复杂了，遂采用上面的三条。
- 令$E_1 = S, E_2 = \cdots = \emptyset$，得到$P(\emptyset) = 0$

## 「一些简单的命题」

$$
\begin{gathered}
prop1: P(E^c) = 1 - P(E) \\
prop2: If\ E \subset F, then\ P(E) \leq P(F) \\
prop3: P(E \cup F) = P(E) + P(F) - P(EF) \\
prop4: P(\cup_{i=1}^{n} E_i) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1 < i_2 < \cdots < i_r} P(E_{i_1} \cdots E_{i_r}) \\
Boole's\ inequality: P(\cup_{i=1}^{n} E_i) \leq \sum_{i=1}^{n} P(E_i)
\end{gathered}
$$

- 证明思路：
  - 1略
  - 2：$F = E \cup E^c F$
  - 3：$E \cup F = E \cup E^c F$，$F = EF \cup E^c F$
  - 用上三条公理，尤其是第三条

## 「等可能结果的样本空间」

样本空间是有限集，其中的每个结果的发生都是等可能的。则事件E发生的概率为：

$$
P(E) = \frac{E\text{中包含的结果数}}{S\text{中包含的总结果数}}
$$

典例：选帽子
有N个男人开party，将各自的帽子拿出洗乱，求最后没有人拿到他的帽子的概率？
记$E_i$为第i个男人拿到了他的帽子。帽子洗乱后与男人的对应顺序共有N!种结果，其中事件$E_{i_1} E_{i_2} \cdots E_{i_n}$对应(N-n)!种结果
故：$P(E_{i_1} E_{i_2} \cdots E_{i_n}) = \frac{(N-n)!}{N!}$，n个男人可以为N中的任n个，共$\binom{N}{n}$种可能。对n，总概率为$\binom{N}{n} \frac{(N-n)!}{N!} = \frac{1}{n!}$
故：$1 - P(\cup_{i=1}^{N} E_i) = 1 - \sum_{r=1}^{N} (-1)^{r+1} \sum_{i_1 < i_2 < \cdots < i_r} P(E_{i_1} \cdots E_{i_r}) = \sum_{r=0}^{N} (-1)^r \frac{1}{r!} \to e.$

# 条件概率和独立性

## 「条件概率」

$If\ P(F) > 0\ ,\ then\ P(E|F) = \frac{P(EF)}{P(F)}$。
$P(EF) = P(E|F)P(F)$
$P(E_1 E_2 \cdots E_n) = P(E_1) P(E_2|E_1) P(E_3|E_1 E_2) \cdots P(E_n|E_1 \cdots E_{n-1})$
对帽子的题，恰好k个人拿到子集帽子的概率是多少？
记E是某个大小为k的子集中每个人都拿到了子集的帽子，G是其余N-k个人都没拿到自己的帽子，$F_i$是大小为k的子集中的第i个人拿到了自己的帽子
$P(E) = P(F_1) P(F_2|F_1) \cdots P(F_k|F_1 \cdots F_{k-1}) = \frac{1}{N} \cdots \frac{1}{N-k+1} = \frac{(N-k)!}{N!}.$
参照上面选帽子的例子：$P(G|E) = \sum_{i=0}^{N-k} (-1)^i \frac{1}{i!}$
则：$P(EG) = \frac{(N-k)!}{N!} \sum_{i=0}^{N-k} (-1)^i \frac{1}{i!}$
故结果为：$P = \binom{N}{k} P(EG) = \frac{1}{k!} \sum_{i=0}^{N-k} (-1)^i \frac{1}{i!} \to \frac{1}{ek!}$

## 「贝叶斯公式」

$$
\begin{aligned}
P(E) &= P(EF) + P(EF^c) \\
&= P(E|F)P(F) + P(E|F^c)P(F^c) \\
&= P(E|F)P(F) + P(E|F^c)[1 - P(F)]
\end{aligned}
$$

$$
\begin{aligned}
P(A|B) &= \frac{P(AB)}{P(B)} \\
&= \frac{P(B|A)P(A)}{P(B)}
\end{aligned}
$$

Law of total probability: 若$F_1 \cdots F_n$是互斥事件即有且仅有一个会发生，事件E的概率可做如下推导：

$$
\begin{gathered}
\cup_{i=1}^{n} F_i = S \\
E = \cup_{i=1}^{n} EF_i\text{，显然}EF_i\text{也是互斥的，用上公理3有} \\
P(E) = \sum_{i=1}^{n} P(EF_i) = \sum_{i=1}^{n} P(E|F_i)P(F_i)
\end{gathered}
$$

贝叶斯公式：

$$
\begin{aligned}
P(F_j|E) &= \frac{P(EF_j)}{P(E)} \\
&= \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}
\end{aligned}
$$

## 「 独立事件」

事件E与F独立，定义为：

$$
P(EF) = P(E)P(F)
$$

事件E、F、G相互独立，定义为：

$$
\begin{gathered}
P(EFG) = P(E)P(F)P(G) \\
P(EF) = P(E)P(F) \\
P(EG) = P(E)P(G) \\
P(FG) = P(F)P(G)
\end{gathered}
$$

可以推出E会独立于F、G组成的任何形式的事件，如：

$$
\begin{gathered}
P[E(F \cup G)] = P(EF \cup EG) \\
= P(EF) + P(EG) - P(EFG) \\
= P(E)P(F) + P(E)P(G) - P(E)P(FG) \\
= P(E)[P(F) + P(G) - P(FG)] \\
= P(E)P(F \cup G)
\end{gathered}
$$

推广：$E_1, \cdots, E_n$相互独立等价于其中的任何子集$E_{i_1}, \cdots, E_{i_r}$都满足：$P(E_{i_1}, \cdots, E_{i_r}) = P(E_{i_1}) \cdots P(E_{i_r})$
有时一次试验可能由一系列的子试验组成，若每次子试验的所有可能结果都一样，则将每次子试验都称为一次trial

## 「 P(·|F)是概率」

首先P(·|F)满足概率的三个公理，故定义Q(E)=P(E|F)可以当普通的概率来用。
由：

$$
\begin{gathered}
Q(E_1 \cup E_2) = Q(E_1) + Q(E_2) - Q(E_1E_2) \\
= P(E_1 \cup E_2|F) = P(E_1|F) + P(E_2|F) - P(E_1E_2|F)
\end{gathered}
$$

还可以推导：

$$Q(E_1) = Q(E_1E_2) + Q(E_1E_2^c) = Q(E_1|E_2)Q(E_2) + Q(E_1|E_2^c)Q(E_2^c)$$

$$Q(E_1|E_2) = \frac{Q(E_1E_2)}{Q(E_2)} = \frac{P(E_1E_2|F)}{P(E_2|F)} = \frac{\frac{P(E_1E_2F)}{P(F)}}{\frac{P(E_2F)}{P(F)}} = P(E_1|E_2F)$$

下式代入上式有：

$$P(E_1|F) = P(E_1|E_2F)P(E_2|F) + P(E_1|E_2^cF)P(E_2^c|F)$$

定义条件独立：$E_1, E_2$在F成立的条件下独立，则为：将独立定义中的P全部换成Q则可

$$P(E_1|E_2F) = P(E_1|F)$$

或者:

$$P(E_1E_2|F) = P(E_1|F)P(E_2|F)$$

例子：Laplace's rule of succession

盒子里有k+1个硬币，第i个硬币在被掷后头朝上的概率为$\frac{i}{k}, i = 0, \cdots, k$.随机选择一个硬币，掷前n次都是头朝上，求掷前n+1次都是头朝上的概率?

记前n次头朝上为$H_n$，则$P(H_{n+1}|H_n) = \frac{P(H_nH_{n+1})}{P(H_n)} = \frac{P(H_{n+1})}{P(H_n)}$

记选到了第i个硬币为事件$C_i$，则：$P(H_n) = \sum_{i=0}^{k} P(H_n|C_i)P(C_i)$

由：$P(H_n|C_i) = (\frac{i}{k})^n, P(C_i) = \frac{1}{k+1}$

全部代入，有：$P(H_{n+1}|H_n) = \frac{\frac{1}{k}\sum_{i=0}^{k}(\frac{i}{k})^{n+1}}{\frac{1}{k}\sum_{i=0}^{k}(\frac{i}{k})^{n}} \to \frac{\int_0^1 x^{n+1}dx}{\int_0^1 x^n dx} = \frac{n+1}{n+2}$

# 随机变量

## 「随机变量」

定义：定义在样本空间上的实值函数称为随机变量

可列取值的随机变量：使用列举法列出概率

分布函数：对于随机变量X，它的分布函数是：$F(X) = P\{X \leq x\}$

是关于x的不减函数。

## 「离散随机变量」

定义：最多能取可数个（可以为无穷个）可能取值的随机变量。

定义pmf：$p(a) = P\{X = a\}$，只在X的可能取值上为非负，其他x值代入均返回0

$\sum_{i=1}^{\infty} p(x_i) = 1$

## 「期望」

$E[X] = \sum_{x:p(x)>0} xp(x)$

指示函数$I_A$的期望：$E[I_A] = P(A)$.

## 「随机变量函数的期望」

Y=g(X)，将Y当作另一个随机变量来看待。

$$E[g(X)] = \sum_i g(x_i)p(x_i)$$

推论：E[aX+b]=aE[X]+b

## 「方差」

$$记\mu = E[X],$$
$$Var(X) = E[(X-\mu)^2],$$
$$Var(X) = E[X^2] - (E[X])^2,$$
$$Var(aX+b) = a^2 Var(X),$$
$$sd(X) = \sqrt{Var(X)}$$

## 「 伯努利及二项分布随机变量」

一次trial，X=1时称为成功，X=0时称为失败，且pmf为：
$p(0) = P\{X=0\} = 1-p, p(1) = P\{X=1\} = p$，其中$0 \le p \le 1$
称X为伯努利随机变量。
若现在有n次独立的trial，每次成功的概率都是p，则X是二项分布随机变量，参数为(n,p)。pmf为：

$$p(i) = \binom{n}{p} p^i (1-p)^{n-i}, i = 0, 1, \cdots, n$$

由二项式定理有：$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{n} \binom{n}{p} p^i (1-p)^{n-i} = [p+1-p]^n = 1$

### 性质

$$X \sim b(n,p), Y \sim b(n-1,p),$$
$$E[X^k] = npE[(Y+1)^{k-1}],$$
$$E[X] = np,$$
$$Var(X) = E[X^2] - (E[X])^2 = np(1-p)$$

命题1：X~b(n,p)，则P{X=k}先单增后单减，最大值在k为小于等于(n+1)p的最大整数时达到。证明用数列的比值法探究单调性则可。

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

### 计算分布函数

$$F(i) = P\{X \le i\} = \sum_{k=0}^{i} \binom{n}{k} p^k (1-p)^{n-k}, i = 0, 1, \cdots, n$$

计算时，用上上面那个P{X=k}和P{X=k-1}间的关系更简便！

## 「 泊松分布随机变量」

给定正数$\lambda$，定义pmf为：$p(i) = P\{X=i\} = e^{-\lambda}\frac{\lambda^i}{i!}, i = 0, 1, 2, \cdots$

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$$

### 使用泊松分布随机变量来估计二项分布

当n非常大，但p很小以至于np是一个适中的数时，令$\lambda = np$，X~b(n,p)
则：

$$P\{X=i\} = \frac{n(n-1)\cdots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^i} \to e^{-\lambda}\frac{\lambda^i}{i!}$$

### 性质

$$E[X]=\sum_{i=0}^{\infty}\frac{ie^{-\lambda}\lambda^i}{i!}=\lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}=\lambda,\ (j=i-1)$$

$$E[X^2]=\sum_{i=0}^{\infty}\frac{i^2e^{-\lambda}\lambda^i}{i!}=\lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{(j+1)\lambda^j}{j!}=\lambda[\lambda+1],\ (j=i-1),$$

$$Var(X)=\lambda[\lambda+1]-\lambda^2=\lambda$$

## 「其他离散分布」

### 几何分布

第1次success发生时trial的次数

$$p(i)=p(1-p)^i,$$
$$E[X]=\frac{1}{p},$$
$$Var(X)=\frac{1-p}{p^2},$$

### 负二项分布

第r次success发生时trial的次数

$$p(i)=\binom{i-1}{r-1}p^r(1-p)^{i-r}, i\geq r,$$
$$E[X]=\frac{r}{p},$$
$$Var(X)=\frac{r(1-p)}{p^2}$$

### 超几何分布

从有m个白球的N个球中随机选取n个球，白球的个数。

$$p(i)=\frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}, i=0,\cdots,m,$$
$$with\ p=\frac{m}{N}, E[X]=np,$$
$$Var(X)=\frac{N-n}{N-1}np(1-p)$$

### 结论

$$E[\sum_{i=1}^{n}X_i]=\sum_{i=1}^{n}E[X_i]$$

## 「累计分布函数的性质」

- F是不减的函数
- $\lim_{b\to\infty}F(b)=1$
- $\lim_{b\to-\infty}F(b)=0$
- F是右连续的，$\lim_{n\to\infty}F(b_n)=F(b)$
- $P\{a<X\leq b\}=F(b)-F(a)$

$$P\{X<b\}=P\{\lim_{n\to\infty}\{X\le b-\frac{1}{n}\}\}$$
$$=\lim_{n\to\infty}P(X\le b-\frac{1}{n})$$
$$=\lim_{n\to\infty}F(b-\frac{1}{n})$$

# 连续型随机变量

## 「 介绍」

离散随机变量是有限个或可数多个取值，而连续型随机变量有不可数多个取值。
对任意实数域的子集B，若存在在$(-\infty,\infty)$上定义的非负可测函数f，满足$P\{X\in B\}=\int_B f(x)dx$，则称f为pdf。
由于X总要取值，则有$1=P\{X\in(-\infty,\infty)\}=\int_{-\infty}^{\infty}f(x)dx$
单点集的概率为0（不可测集上的概率为0）

$$P\{X\le a\}=P\{X<a\}=F(a),$$
$$\frac{d}{da}F(a)=f(a)$$

例：已知X的pdf为$f_X$,求Y=2X的pdf?

$$F_Y(a)=P\{Y\le a\}=P\{2X\le a\}=F_X(\frac{a}{2}),$$
$$\text{求导得，}f_Y(a)=\frac{1}{2}f_X(\frac{a}{2})$$

## 「 期望和方差」

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

定理：

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

证明略，见课本P194~P195
期望的线性运算性质，方差的定义，方差的线性运算性质同离散随机变量

## 「 均匀分布随机变量」

X~U($\alpha,\beta$)，则pdf为：$f(x)=\frac{1}{\beta-\alpha},\alpha<x<\beta,0,otherwise$

$$E[X]=\frac{\beta+\alpha}{2},$$
$$Var(X)=\frac{(\beta-\alpha)^2}{12}$$

## 「 正态分布随机变量」

$$X\sim N(\mu,\sigma),$$
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<\infty,$$
$$E[X]=\mu,$$
$$Var(X)=\sigma^2$$

要证pdf在支撑集上积分为1，要求$I=\int_{-\infty}^{\infty}e^{-\frac{y^2}{2}}dy$，用极坐标代换求$I^2$则可。
标准化：

$$
\begin{gathered}
Z=\frac{X-\mu}{\sigma}, \\
F_Z(x)=P\{Z\le x\} \\
=P\{X\le x\sigma+\mu\} \\
=F_X(x\sigma+\mu), \\
\text{记}F_Z(x)=\phi(x), \\
f_Z(x)=\sigma f_X(x\sigma+\mu)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},-\infty<x<\infty, \\
Z\sim N(0,1), \\
\text{可用积分}I\text{与分部积分求得，}E[Z]=0,Var(Z)=1, \\
\text{代回}X\text{，可得}E[X],Var(X)
\end{gathered}
$$

### 用正态分布近似二项分布

是中心极限定理的一个特例。

$$
\begin{gathered}
S_n\sim b(n,p), \\
P\{a\le\frac{S_n-np}{\sqrt{np(1-p)}}\le b\}\to\phi(b)-\phi(a),n\to\infty
\end{gathered}
$$

## 「 指数分布随机变量」

$$
\begin{gathered}
X\sim exp(\lambda), \\
f(x)=\lambda e^{-\lambda x},x\ge 0,0,otherwise, \\
F(a)=1-e^{-\lambda a},a\ge 0,0,ohterwise, \\
E[X^n]=\frac{n}{\lambda}E[X^{n-1}], \\
E[X]=\frac{1}{\lambda}, \\
Var(X)=\frac{1}{\lambda^2}
\end{gathered}
$$

**无记忆性**: $P\{X>s+t|X>t\}=P\{X>s\}$
危险率函数，见生存分析

## 「 其他连续型分布」

见数理统计复习

## 「 随机变量的函数的分布」

### 求cdf关系再求导

$$
\begin{gathered}
Y=g(X), \\
F_Y(x)=P\{Y\le x\} \\
=P\{g(X)\le x\} \\
=P\{X\le g^{-1}(x)\} \\
=F_X(g^{-1}(x)), \\
\text{再两边求导可以得到}pdf
\end{gathered}
$$

例子: $Y=X^n,Y=|X|$

**总结为：**

$$Y=g(X),$$
$$f_Y(y)=f_X[g^{-1}(y)]|\frac{d}{dy}g^{-1}(y)|,y=g(x)\text{时}$$

# 随机变量的联合分布

## 「 联合分布函数」

$$continuous\ rv$$
$$F(a,b)=P\{X\le a,Y\le b\},$$
$$P\{a_1\le X\le a_2,b_1\le Y\le b_2\}=F(a_2,b_2)+F(a_1,b_1)-F(a_1,b_2)-F(a_2,b_1),$$
$$F_X(a)=P\{X\le a,-\infty\le y\le\infty\}=F(a,\infty),$$
$$F_Y(b)=F(\infty,b)$$
$$discrete\ rv$$
$$p(x,y)=P\{X=x,Y=y\},$$
$$p_X(x)=P\{X=x\}=\sum_j p(x,y_j),$$
$$p_Y(y)=\sum_i p(x_i,y)$$

定义pdf：

$$continuous\ rv$$
$$\text{重要：}P\{(X,Y)\in C\}=\iint_{(X,Y)\in C}f(x,y)dxdy\text{，用重积分的方法进行计算（在已知联合}pdf\text{时）}$$
$$F(a,b)=\int_{-\infty}^{b}\int_{-\infty}^{a}f(x,y)dxdy,$$
$$f(a,b)=\frac{\partial^2}{\partial a\partial b}F(a,b),$$
$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dx,$$
$$P\{X\in A\}=\int_A f_X(x)dx$$

## 「 独立随机变量」

$$X,Y\ are\ independent\text{，等价于}$$
$$\text{任取}A,B\text{，均有}P\{X\in A,Y\in B\}=P\{X\in A\}P\{Y\in B\},$$
$$continuous\ rv,F(a,b)=F_X(a)F_Y(b),\text{也即}f(x,y)=f_X(x)f_Y(y),$$
$$discrete\ rv,p(x,y)=p_X(x)p_Y(y)$$

**典例**：一天进入邮局的人数服从$Pois(\lambda)$，证明：每个进入邮局的人为男性的概率为p，为女性的概率为1-p，则进入邮局的男性个数X和进入邮局的女性个数Y是独立的泊松变量，且参数分别为$\lambda p,\lambda(1-p)$.

$$\text{只需求联合密度则可，}$$
$$P\{X=x,Y=y\}=P\{X=x,Y=y|X+Y=i+j\}P\{X+Y=i+j\}+$$
$$P\{X=x,Y=y|X+Y\ne i+j\}P\{X+Y\ne i+j\}$$
$$=P\{X=x,Y=y|X+Y=i+j\}P\{X+Y=i+j\}$$
$$=\binom{i+j}{i}p^i(1-p)^je^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}\text{，再用这个式子分别求}P\{X=i\}\text{和}P\{Y=y\}\text{则可}$$

中间有个条件概率是二项分布的概率

**典例**：X~U(0,60), Y~U(0,60), X与Y相互独立，求P{|X-Y|<10}

$$\begin{aligned}P\{|X-Y|>10\} &= P\{X+10<Y\}+P\{Y+10<X\}\\ &= 2P\{X+10<Y\}\\ &= 2\iint_{x+10<y} f_X(x)f_Y(y)dxdy\\ &= 2\int_{10}^{60}\int_0^{y-10}(\frac{1}{60})^2dxdy=\frac{25}{36},\\ P\{|X-Y|<10\} &= 1-P\{|X-Y|>10\}=\frac{11}{36}\end{aligned}$$

**典例**：期中考试。$f(x,y)=\frac{1}{2},0\le X\le 2,0\le Y\le 1$，求$P\{XY\le a\},(a>0)$

$$\begin{aligned}P\{XY\le a\} &= \iint_{0\le x\le\frac{a}{y}\le 2}\frac{1}{2}dxdy\\ &= \int_{\frac{a}{2}}^1\int_0^{\frac{a}{y}}\frac{1}{2}dxdy+\int_0^{\frac{a}{2}}\int_0^2\frac{1}{2}dxdy\\ &= \frac{a}{2}\ln\frac{2}{a}+\frac{a}{2}\end{aligned}$$

怎么做重积分？在平面直角坐标系中画出积分区域，再看先积哪个轴。如上题，就是在坐标系中画出了区域，其中$XY\le a$这个区域是双曲线往原点的一侧，代入(0,0)可见！（上题是对区域先积x轴，做了一条平行于x轴的直线）

**典例**：布芬投针实验。无限长的桌子宽为D，针长为L≤D，计算针与其中一条边相交的概率？
记针的中点到最近边的距离为X~U(0, $\frac{D}{2}$)，针所在直线与两条边的夹角的余角为$\theta\sim U(0,\frac{\pi}{2})$，题中事件即为：$\{\frac{X}{\cos\theta}<\frac{L}{2}\}$.

$$P\{X<\frac{L}{2}\cos\theta\}=\iint_{x<\frac{L}{2}\cos y}f_X(x)f_\theta(y)dxdy=\frac{4}{\pi D}\int_0^{\frac{\pi}{2}}\int_0^{\frac{L}{2}\cos y}dxdy=\frac{2L}{\pi D}$$

画出积分区域，结果是一个很简单的区域，因为有$L\le D$这个条件。

又例：区域$yz\le x\le 1,0\le y\le 1,0\le z\le 1$，直接写则可，因为对任何$0\le y\le 1,0\le z\le 1$，都满足了$yz\le 1$

X与Y独立，当且仅当联合pdf（或联合pmf）可以如下分解：

$$f_{X,Y}(x,y)=h(x)g(y)$$

例子：$X_1,X_2,\cdots$是一列iid的随机变量，记事件$A_n$为$X_n>X_i,i=1,\cdots,n-1$，问$A_{n+1}$与$A_n$是否独立？
反过来考虑，显然$P(A_n|A_{n+1})=P(A_n)=\frac{1}{n}$，则独立。

## 「 独立随机变量的和」

卷积公式：

$$\begin{aligned}F_{X+Y}(a) &= \int_{-\infty}^{\infty}F_X(a-y)f_Y(y)dy,\\ f_{X+Y}(a) &= \int_{-\infty}^{\infty}f_X(a-y)f_Y(y)dy\end{aligned}$$

a-y应落在X的支撑集中，据此得到y和a的约束关系，在y~a坐标系中进行积分。
**典例**：X~U(0,1), Y~U(0,1), X与Y相互独立，求X+Y的pdf?
法一：用上节的方法

$$F(X+Y\le a)=\iint_{X+Y\le a,0\le X,Y\le 1}dxdy=\frac{1}{2}a^2,a\le 1,2a-\frac{1}{2}a^2,1<a<2,\\ \text{求导得到}pdf$$

法二：

$$f_{X+Y}(a)=\int_0^1 f_X(a-y)dy,$$
$$\text{要满足：}0\le a-y\le 1, 0\le y\le 1, \text{画出}y\sim a\text{图，由图讨论并分段积分，}$$
$$\text{得到一样的结果}$$

### 正态随机变量

正态分布可加性，自然参数对应相加则可。求法类比上例。

$$N(\mu_1,\sigma_1^2)+N(\mu_2,\sigma_2^2)=N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$$

### 泊松和二项分布随机变量

$$P\{X+Y=n\}=\sum_{k=0}^{n}P\{X=k,Y=n-k\},\dots$$

也有可加性：

$$b(n_1,p)+b(n_2,p)=b(n_1+n_2,p),$$
$$pois(\lambda_1)+pois(\lambda_2)=pois(\lambda_1+\lambda_2)$$

## 「 离散情形下的条件分布」

$$\text{给定}Y=y\text{时}X\text{的条件分布：}$$
$$p_{X|Y}(x|y)=P\{X=x|Y=y\}$$
$$=\frac{P\{X=x,Y=y\}}{P\{Y=y\}},$$
$$F_{X|Y}(x|y)=P\{X\le x|Y=y\}=\sum_{a\le x}p_{X|Y}(a|y)$$

## 「 连续情形下的条件分布」

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)},$$
$$P\{X\in A|Y=y\}=\int_A f_{X|Y}(x|\frac{f(x,y)}{f_Y(y)}=y)dx,$$
$$\text{当}X\text{，}Y\text{独立时，有}f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{f_X(x)f_Y(y)}{f_Y(y)}=f_X(x)$$

**典例**：求T分布的pdf?

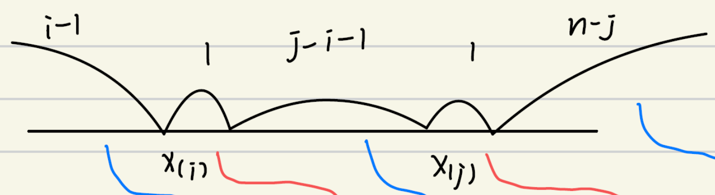
$$T=\frac{Z}{\sqrt{\frac{Y}{n}}}=\sqrt{\frac{n}{Y}}Z,$$
$$\text{先算}f_{T|Y}(t|y),\text{即为}N(0,\frac{n}{Y})\text{的}pdf,$$
$$f_{T,Y}(t,y)=f_{X|Y}(x|y)f_Y(y)\text{可求出}f_{T,Y}(t,y),$$
$$f_T(t)=\int_0^\infty f_{T,Y}(t,y)dy$$

连续given离散：

$$f_{X|N}(x|n)=\lim_{dx\to 0}\frac{P\{x<X<x+dx|N=n\}}{dx}=\lim_{dx\to 0}\frac{P\{N=n|x<X<x+dx\}P\{x<X<x+dx\}}{P\{N=n\}dx}$$
$$=\frac{P\{N=n|X=x\}}{P\{N=n\}}f(x)$$

## 「 次序统计量」

记为：$X_{(1)}, X_{(2)}, \cdots, X_{(n)}$




$$f(x_i, x_j) = \binom{n}{i-1,\ 1,\ j-i-1,\ 1,\ n-j} [F(x_i)]^{i-1} f(x_i) [F(x_j) - F(x_i)]^{j-i-1} f(x_j) [1 - F(x_j)]^{n-j}$$

$$\binom{n}{i-1,\ 1,\ j-i-1,\ 1,\ n-j} = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}$$

## 「 随机变量函数的联合分布」

$$Y_1 = g_1(X_1, X_2), Y_2 = g_2(X_1, X_2), \text{解出旧变量，得到：}$$
$$X_1 = h_1(Y_1, Y_2), X_2 = h_2(Y_1, Y_2),$$
$$J(y_1, y_2) = \frac{\partial h_1}{\partial y_1}\frac{\partial h_2}{\partial y_2} - \frac{\partial h_1}{\partial y_2}\frac{\partial h_2}{\partial y_1} = \frac{D(X_1, X_2)}{D(Y_1, Y_2)}, \text{新变量在下}$$
$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(h_1(Y_1, Y_2), h_2(Y_1, Y_2))|J(y_1, y_2)|, \text{注意新变量的取值范围}$$

**典例**：期中考试。$f(x, y) = \frac{1}{2}, 0 \le X \le 2, 0 \le Y \le 1$，求$P\{XY \le a\}, (a > 0)$

$$\text{令} Z = XY, V = Y,$$
$$J = \frac{1}{2},$$
$$f_{Z,V}(z, v) = \frac{1}{2v}, 0 < v < 1, 0 < \frac{z}{v} < 2,$$
$$\text{画个图则可，是一个} z \sim v \text{平面中的三角形区域，这里先积} v\text{，自然} z \text{也会有个范围，}$$
$$f_Z(z) = \int_{\frac{2}{z}}^{1} \frac{1}{2v} dv = \frac{1}{2} \ln\left(\frac{2}{z}\right), 0 < z < 2$$

# 期望的性质

## 「 介绍」

如果$P\{a \le X \le b\} = 1$，则$a \le E[X] \le b$

$$E[X] = \sum_{x:p(x)>0} xp(x) \ge \sum_{x:p(x)>0} ap(x) = a \sum_{x:p(x)>0} p(x) = a$$

## 「 随机变量和的期望」

$$E[g(X, Y)] = \sum_y \sum_x g(x, y) f(x, y) dx dy,$$
$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$
$$E[X + Y] = E[X] + E[Y]$$

利用指示变量的期望性质证明不等式：（凑事件，又见书315）

$$X_i = 1, if\ A_i\ occurs, 0, otherwise, X = \sum_{i=1}^{n} X_i,$$
$$Y = 1, if\ X \geq 1, 0, otherwise,$$
$$\text{则有：} X \geq Y, E[X] \geq Y,$$
$$\text{得到：} P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$$

典例：邮票收集问题
一共有N种邮票，每次等可能地得到一种邮票，X为收集满所有类型邮票为止时收集的邮票数，求E[X]?

$$\text{令}X_i\text{为收集了}i\text{种邮票后直到获得新种类的邮票为止需要再收集的邮票数}, i = 0, \cdots, N-1, \text{显然}X_0 = 1$$
$$\text{有}X = X_0 + X_1 + \cdots + X_{N-1} P\{X_i = k\} = (\frac{i}{N})^{k-1} \frac{N-i}{N} k \leq 1, \text{得：} E[X_i] = \frac{N}{N-i},$$
$$E[X] = N(1 + \cdots + \frac{1}{N})$$

有用的恒等式：

$$\text{对任意非负整数随机变量}X\text{，令}X_i = 1, if\ X \geq i, 0, otherwise,$$
$$\sum_{i=1}^{\infty} X_i = X,$$
$$E[X] = \sum_{i=1}^{\infty} E[X_i] = \sum_{i=1}^{\infty} P\{X \geq i\},$$
$$\text{连续情形同理，则为：} E[X] = \int_0^{\infty} P(X > y) dy$$

## 「 试验序列中发生次数的矩」

$$I_i = 1, \text{如果}A_i\text{出现}, 0, otherwise,$$
$$X = \sum_{i=1}^{n} I_i, \binom{X}{2} = \sum_{i<j} I_i I_j,$$
$$E[\binom{X}{2}] = \sum_{i<j} E[I_i I_j] = \sum_{i<j} P(A_i A_j),$$
$$\text{则}E[\frac{X(X-1)}{2}] = \sum_{i<j} P(A_i A_j), \text{得到}E[X^2]$$

同理可以得到更高阶矩$\binom{X}{k} = \sum_{i_1<i_2<\cdots<i_k} I_{i_1} I_{i_2} \cdots I_{i_k}$
例子：邮票收集问题2
共N种不同的邮票，设第j种邮票出现的概率为$p_j$，令Y为前n个邮票中收集的邮票种类数，求E[Y], Var(Y)?

$$X = N - Y\text{为还没收集到的种类数，令}A_i\text{为第}i\text{种邮票还没出现，则}X\text{为}A_1, \cdots, A_N\text{中发生的数量},$$
$$P(A_i) = (1-p_i)^n, E[X] = \sum_{i=1}^{N} (1-p_i)^n,$$
$$P(A_i A_j) = (1 - p_i - p_j)^n, i \neq j, \text{代入上面过程可求出}E[X^2]$$

## 「 随机变量和的协方差、方差及相关系数」

$$\text{如果}X\text{、}Y\text{相互独立，则}E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$
$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

$$Cov(X,Y) = Cov(Y,X),$$
$$Cov(X,X) = Var(X),$$
$$Cov(aX,Y) = aCov(X,Y),$$
$$Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i,Y_j)$$

$$Var(\sum_{i=1}^{n} X_i) = Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} X_j)$$
$$= \sum_{i=1}^{n} Var(X_i) + \sum_{i\neq j} Cov(X_i,X_j)$$
$$= \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i,X_j)$$

$$X = X_1 + \cdots + X_n,$$
$$Var(X_i) = E[X_i^2] - (E[X_i])^2 = E[X_i] - (E[X_i])^2 = p - p^2 Var(X) = Var(X_1) + \cdots + Var(X_n)$$

相关系数:

$$-1 \le \rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \le 1$$

判定相关系数的符号:

$$Cov(I_A, I_B) = P(AB) - P(A)P(B) = P(B)[P(A|B) - P(A)]\text{，取决于}P(A|B)\text{与}P(A)\text{的相对大小}$$

## 「 条件期望」

$$E[X|Y=y] = \sum_x xP\{X|Y\}$$

例子：X,Y均服从b(n,p)，求X给定X+Y=m时的条件期望？
要求P{X=k|X+Y=m}，分子用X，Y的pdf可求出，分母如下：
将X+Y当作新变量，服从b(2n,p)，据此可以求分母。
发现上面求出的这个概率是超几何分布的pdf

### 重期望

$$E[X] = E[E[X|Y]]$$

即证明：$E[X] = \sum_y E[X|Y=y]P\{Y=y\}$（因为E[X|Y]是y的函数），打开则可。

### 用条件期望求概率或期望（好用！）

重要思路：用条件期望将一个随机变量**非随机化再随机化**。如下例子：

$$E[XY] = E[E[XY|Y]],$$
$$E[XY|Y=y] = E[Xy|Y=y] = yE[X|Y=y] = yE[X](\text{如果}X\text{，}Y\text{相互独立的话，下同}),\text{非随机化}$$
$$E[XY|Y] = YE[X],\text{再随机化},$$
$$E[XY] = E[YE[X]] = E[X]E[Y]\text{（如果}X\text{，}Y\text{相互独立的话）}$$

重要思路：设X是事件A的指示变量，则
$E[X] = P(A) = E[E[X|Y]] = \sum_y E[X|Y=y]P\{Y=y\} = \sum_y P(A|Y=y)P\{Y=y\}$
Y连续时，为：$P(A) = \int_{-\infty}^{\infty} P\{A|Y=y\}f_Y(y)dy$

例子：X和Y都是独立连续的随机变量，pdf分别为$f_X, f_Y$，计算P{X<Y}。

$$
\begin{aligned}
P\{X<Y\} &= \int_{-\infty}^{\infty} P\{X<Y|Y=y\} f_Y(y)dy \\
&= \int_{-\infty}^{\infty} P\{X<y|Y=y\} f_Y(y)dy, \\
&= \int_{-\infty}^{\infty} P\{X<y\} f_Y(y)dy, \\
&= \int_{-\infty}^{\infty} F_X(y) f_Y(y)dy
\end{aligned}
$$

### 条件方差（厉害）

$$
Var(X|Y) = E[(X-E[X|Y])^2|Y],
$$
$$
Var(X|Y) = E[X^2|Y] - (E[X|Y])^2
$$

对上式两边取期望，再结合Var(E[X|Y])套Var(X)=E[X^2]-(E[X]) ^2的结果，可以得到：

$$
Var(X) = E[Var(X|Y)] + Var(E[X|Y])
$$

往往可用上个小节：非随机化->随机化的思想来求右边两项，从而得到Var(X)

例子：$X_1, X_2, \cdots$是一列iid的随机变量，N是一个非负整数随机变量且与$X_i, i \geq 1$独立，求$Var(\sum_{i=1}^{N} X_i)$?

$$
E[\sum_{i=1}^{N} X_i|N] = NE[X], (\text{按照非随机化}->\text{随机化来操作就行})
$$

$$
Var(\sum_{i=1}^{N} X_i|N) = N\,Var(X),
$$

$$
Var(\sum_{i=1}^{N} X_i) = E[N]Var(X) + (E[X])^2 Var(N)
$$

## 「 条件期望和预测」

$$
\text{模型：} Y = g(X) + e, \text{当用}MSE\text{准则时，}
$$
$$
E[(Y-g(X))^2] \geq E[(Y-E[Y|X])^2]
$$