第三章 随机变量与分布函数

随机变量及其分布

随机变量

设$\xi(w)$是定义在概率空间$(\Omega,\mathcal{F},P)$上的单值实函数，如果对直线上任一$Borel$点集$\mathbf{B}$，有$\{w:\xi(x)\in\mathbf{B}\}\in\mathcal{F}$，则称$\xi(w)$为随机变量，而$P\{\xi(w)\in\mathbf{B}\}$称为随机变量的概率分布

$P(A)$一定要用$P\{\xi<x\}$来表示，不能直接写$P(\xi)$

分布函数

- 称$F(x)=P\{\xi(w)<x\},-\infty<x<+\infty$为随机变量$\xi(w)$的分布函数
- 将"随机变量$\xi(w)$服从分布函数$F(x)$"记作$\xi(w)\sim F(x)$
- $P\{a\leqslant\xi(w)<b\}=F(b)-F(a)$
- 性质
  - 单调性：若$a<b$，则$F(a)<F(b)$
  - $\lim\limits_{x\to-\infty}F(x)=0,\ \lim\limits_{x\to+\infty}F(x)=1$

    由于$P\{-\infty<x<+\infty\}=\sum\limits_{n=-\infty}^{+\infty}P\{n\leqslant x<n+1\}$
    $=\sum\limits_{n=-\infty}^{+\infty}[F(n+1)-F(n)]$
    $=\lim\limits_{n\to+\infty}F(n)-\lim\limits_{m\to-\infty}F(m)=1$
    且由于$F(x)$单调有界，$\lim\limits_{x\to-\infty}F(x)=\lim\limits_{m\to-\infty}F(m);\ \lim\limits_{x\to+\infty}F(x)=\lim\limits_{n\to+\infty}F(n)$
    即证明极限的存在性
    因为$0\leqslant F(x)\leqslant1$，故$\lim\limits_{x\to-\infty}F(x)=0,\ \lim\limits_{x\to+\infty}F(x)=1$
  - 左连续性：$F(x-0)=F(x)$

    由于$F(x)$单调，只需证明对于一列单调递增趋于x的数列$\{x_n\}$，也有$\lim\limits_{n\to\infty}F(x_n)=F(x)$
    $F(x)-F(x_0)=P\{x_0\leqslant\xi<x\}$
    $=\sum\limits_{n=0}^{\infty}P\{x_n\leqslant\xi<x_{n+1}\}$
    $=\sum\limits_{n=0}^{\infty}[F(x_{n+1})-F(x_n)]$
    $=\lim\limits_{n\to\infty}F(x_n)-F(x_0)$
    即有$F(x-0)=\lim\limits_{n\to\infty}F(x_n)=F(x)$

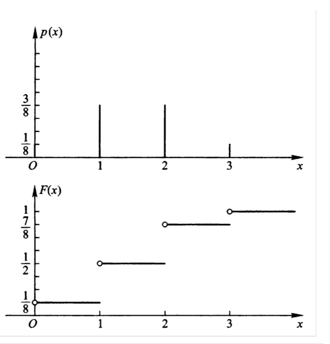
  满足这3个条件的$F(x)$一定是某个随机变量的分布函数
- 有了分布函数，关于随机变量$\xi(w)$的许多概率都可以方便的算出来，例如
  $P\{\xi(w)=a\}=F(a+0)-F(a)$ 若连续，则单点概率是0
  $P\{\xi(w)\leqslant a\}=F(a+0)$
  $P\{\xi(w)\geqslant a\}=1-F(a)$
  $P\{\xi(w)>a\}=1-F(a+0)$

离散型随机变量

- 设$\{x_i\}$为离散型随机变量ξ的所有可能值，而$p(x_i)$是ξ取x_i的概率，即$P\{\xi=x_i\}=p(x_i),i=1,2,3\ldots$
  $\{p(x_i),i=1,2,3\ldots\}$称为随机变量ξ的概率分布，它满足：$p(x_i)\geqslant0,i=1,2,3\ldots,\sum\limits_{i=1}^{\infty}p(x_i)=1$
- 常用分布列来表示离散型随机变量ξ的概率分布

ξ	x_1	x_2	$\cdots$	x_n	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

- 有了分布列，就可用$F(x)=P\{\xi<x\}=\sum\limits_{x_k<x}p(x_k)$来求得分布函数$F(x)$(跳跃函数)



离散型随机变量及其概率分布的例子

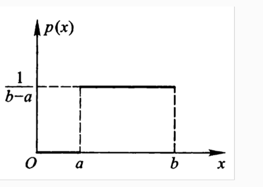
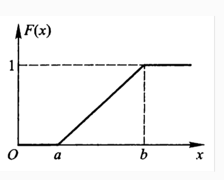
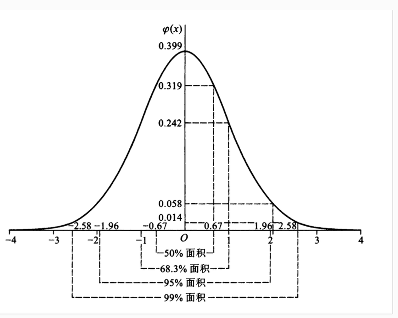
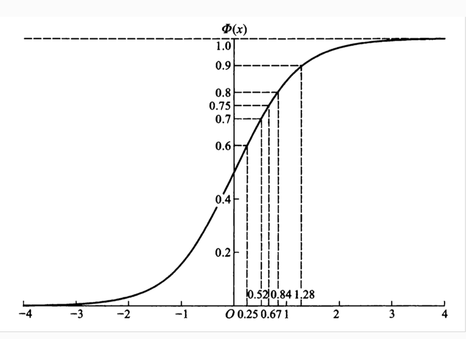
	意义	概率分布
[退化分布/单点分布]	随机变量α只取常数c	$P\{\alpha=c\}=1$
[伯努利分布/两点分布] (n=1时是二项分布)	一次试验中，事件A出现的概率为p，不出现的概率为$q=1-p$，用β记事件A出现的次数，则β只取0，1两值	$b_k=P\{\beta=k\}=p^kq^{1-k},k=0,1$
[二项分布] $\mu\sim B(n,p)$	n重伯努利试验中，若以μ记成功的次数，则它是一个随机变量，μ可能的取值为$0,1,2,\ldots,n$	$b(k;n,p)=P\{\mu=k\}=\binom{n}{k}p^kq^{n-k},k=0,1,2\ldots,n$
[超几何分布]	N件产品有M件次品，不放回抽样调查n件产品，出现次品数ν是随机变量，取值为$0,1,2,\ldots,n$	$h_k=P\{\nu=k\}=\dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{k}},0\leqslant k\leqslant n\leqslant N,k\leqslant M$
[Poisson分布] $\xi\sim P(\lambda)$	随机变量ξ可取到一切非负整数值	$P\{\xi=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda},k=0,1,2,\ldots$，其中$\lambda>0$
[几何分布] $\xi\sim g(k,p)$	在成功的概率为p的伯努利试验中，以η记首次出现成功时的试验次数，则η是随机变量，取值为$1,2,3,\ldots$	$g(k,p)=P\{\eta=k\}=q^{k-1}p,k=0,1,2\ldots$
[帕斯卡分布] (r=1时是几何分布)	在成功的概率为p的伯努利试验中，以ζ记第r次成功时的试验次数，则ζ是随机变量，取值为$r,r+1,\ldots$	$P\{\zeta=k\}=\binom{k-1}{r-1}p^rq^{k-r},k=r,r+1,\ldots$
[负二项分布]		$Nb(l;n,p)=\binom{-r}{l}p^r(-q)^l,l=0,1,2\ldots$

- 几何分布的无记忆性
  - 假设前m次试验没有出现成功，为了达到首次成功还需要等待的时间为η'，
    其概率分布为$P\{\eta'=k\}=P\{\eta=k+m|\eta>m\}$
    $=\dfrac{P\{\eta=k+m\}}{P\{\eta>m\}}$
    $=\dfrac{q^{m+k-1}p}{q^m}$
    $=q^{k-1}p,\quad k=0,1,2\ldots$
    还是服从几何分布，与前面的失败次数m无关
  - 在离散分布中，只有几何分布"永远年青"，即：若η是取正整数的随机变量，且已知$\eta>k$的条件下，$\eta=k+1$的概率与k无关，那么η服从几何分布
    pf：设$p=P\{\eta=k+1|\eta>k\}$，记$q_k=P\{\eta>k\},p_k=\{\eta=k\}$，那么$p_{k+1}=q_k-q_{k+1}$
    则$p=\dfrac{p_{k+1}}{q_k}=\dfrac{q_k-q_{k+1}}{q_k}=1-\dfrac{q_{k+1}}{q_k}$，即$\dfrac{q_{k+1}}{q_k}-=1-p$
    由$q_0=1,q_k=(1-p)^k$，因此$p_k=(1-p)^{k-1}-(1-p)^k=p(1-p)^{k-1},k=1,2,\ldots$

连续型随机变量

- 随机变量ξ可取某个区间$[c,d]$或$(-\infty,+\infty)$中的一切值，且其分布函数$F(x)$是绝对连续函数
- 密度函数
  - $\exists$可积的函数$p(x)$，$s.t.\ F(x)=\int_{-\infty}^{x}p(y)\mathrm{d}y$
  - 由分布函数的性质可知，对应有$\begin{cases}p(x)\geqslant0\\ \int_{-\infty}^{+\infty}p(x)\mathrm{d}x=1\end{cases}$
    反之，若$p(x)$是定义在$(-\infty,+\infty)$上的可积函数且具有如上两条性质，则定义的$F(x)$也满足分布函数的三条性质
  - $p(x)\Delta x\approx\int_{x}^{x+\Delta x}p(y)\mathrm{d}y=F(x+\Delta x)-F(x)$
    密度函数$p(x)$反映了随机变量取x的临近值的概率大小
- 连续型随机变量取个别值的概率为0
- 性质
  - $p(x)=F'(x)$
  - $P(\xi=x)=0,\forall x\in\mathbb{R}$
  - $P(\xi\in[a,b])=\int_a^b p(x)\mathrm{d}x$
  - $P(\xi\geqslant x)=\int_x^{\infty}p(y)\mathrm{d}y$

连续型分布的例子

	密度函数	分布函数
[均匀分布] $\xi\sim U[a,b]$	$p(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & x<a或x>b\end{cases}$ 	$F(x)=\begin{cases}0, & x\leqslant a\\ \dfrac{x-a}{b-a}, & a<x\leqslant b\\ 1, & x>b\end{cases}$ 
[正态分布] $\xi\sim N(\mu,\sigma^2)$	$p(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty$ 当$\mu=0,\sigma=1$时，分布称为标准正态分布 记为$N(0,1)$，分布密度函数记为$\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 	$F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(y-\mu)^2}{2\sigma^2}}\mathrm{d}y,-\infty<x<+\infty$ 标准正态分布的分布函数记为$\Phi(x)=\dfrac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{y^2}{2}}\mathrm{d}y$ 
[指数分布] $\xi\sim\mathrm{Exp}(\lambda)$	$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant0\\ 0, & x<0\end{cases}$	$F(x)=\begin{cases}1-e^{-\lambda x}, & x\geqslant0\\ 0, & x<0\end{cases}$
[Erlang分布]	$f(x)=\begin{cases}0 & ,x\leqslant0\\ \dfrac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x} & ,x>0\end{cases}$	
[Γ分布] $\xi\sim G(\lambda,r)$	$f(x)=\begin{cases}0 & ,x\leqslant0\\ \dfrac{\lambda^n}{(n-1)!}x^{n-1}e^{-\lambda x} & ,x>0\end{cases}$，$\lambda>0,r>0$为常数	

- 指数分布的无记忆性
  - $\forall s>0,t>0,P\{\xi\geqslant s+t|\xi\geqslant s\}=\dfrac{P\{\xi\geqslant s+t\}}{P\{\xi\geqslant s\}}=\dfrac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=P\{\xi\geqslant t\}$
  - 在连续分布中，只有指数分布"永远年青"，即：若$\xi\geqslant0$，其分布函数为$F(x)$
    记$G(x)=P\{\xi\geqslant x\}$，由于$P\{\xi\geqslant s+t|\xi\geqslant s\}=P\{\xi\geqslant t\}$，得$G(s+t)=G(s)G(t)$，$\forall s,t>0$
    又由$G(x)$单调，则$\exists a\in(0,1)$，$s.t.\ G(x)=a^x,x\geqslant0$，令$\lambda=-\ln a>0$，即$a=e^{-\lambda}$
    则$F(x)=1-G(x)=1-e^{-\lambda}$，即为指数分布

随机向量，随机变量的独立性

- 一个二元函数$F(x,y)$是一个二维随机向量的分布函数 $\iff$ $\begin{cases}(i)F(x,y)是关于每个变元单调不减的\\(ii)F(-\infty,y)=0,F(x,-\infty)=0,F(+\infty,+\infty)=1,\forall x,y\in\mathbb{R}\\(iii)F(x,y)关于每个变量均是左连续的\\(iv)P(\xi\in[a_1,b_1),\eta\in[a_2,b_2))=F(b_1,b_2)-F(a_1,b_2)-F(a_2,b_1)+F(a_1,a_2)\geqslant0\end{cases}$
- $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}p(u,v)\mathrm{d}u\mathrm{d}v$
  $F_1(x)=\int_{-\infty}^{x}p_1(u)\mathrm{d}u$
  $F_2(x)=\int_{-\infty}^{y}p_2(v)\mathrm{d}v$
- 条件分布
  - 离散情况：$P\{\eta=y_j|\xi=x_i\}=\dfrac{P\{\xi=x_i,\eta=y_i\}}{P\{\xi=x_i\}}=\dfrac{p(x_i,y_i)}{p_1(x_i)}$
  - 一般情况

    一定要用极限形式来写

    $P\{\eta<y|\xi=x\}=\lim\limits_{\Delta x\to0}P\{\eta<y|x\leqslant\xi<x+\Delta x\}$
    $=\lim\limits_{\Delta x\to0}\dfrac{P\{\eta<y|x\leqslant\xi<x+\Delta x,\eta<y\}}{P\{\eta<y|x\leqslant\xi<x+\Delta x\}}$
    $=\lim\limits_{\Delta x\to0}\dfrac{F(x+\Delta x,y)-F(x,y)}{F(x+\Delta x,\infty)-F(x,\infty)}$
  - 连续情况
    - $P\{\eta<y|\xi=x\}=\int_{-\infty}^{y}\dfrac{p(x,v)}{p_1(x)}\mathrm{d}v$
    - 在给定$\xi=x$的情况下，ξ的分布密度函数为$p(y|x)=\dfrac{p(x,v)}{p_1(x)}$
- 随机变量的独立性
  - 定义
    - 若$\forall x_1,\cdots,x_n$，成立$P\{\xi_1<x_1,\cdots,\xi_n<x_n\}=P\{\xi_1<x_1\}\cdots P\{\xi_n<x_n\}$
    - $\iff F(x_1,\cdots,x_n)=F_1(x_1)\cdots F_n(x_n)$
    - 离散型 $\iff P\{\xi_1=x_1,\cdots,\xi_n=x_n\}=P\{\xi_1=x_1\}\cdots P\{\xi_n=x_n\}$
    - 连续型 $\iff p(x_1,\cdots,x_n)=p_1(x_1)\cdots p_n(x_n)$

    直接套一下定义就行
  - 条件分布化为无条件分布：$P\{\eta<y|\xi=x\}=P\{\eta<y\}$

    证明一定要按一般的情况，用极限的形式来写
  - $\xi_1,\xi_2\sim N(\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho)$，则$\xi_1,\xi_2$独立 $\iff$ $\rho=0$
    $\Rightarrow$当$x=\mu_1,y=\mu_2$时，带入$p(x,y)=p_1(x)p_2(y)$得$\sqrt{1-\rho^2}=1$，即$\rho=0$
    $\Leftarrow$直接验证就可得$p(x,y)=p_1(x)p_2(y)$

随机变量的函数及其分布

- $Borel$函数与随机变量的函数
  - 设$y=g(x)$，若对于$\mathbb{R}^1$中的一切$Borel$点集，均有$\{x:g(x)\in B_1\}\in\mathscr{B}_1$，其中$\mathscr{B}_1$是$\mathbb{R}^1$上的$Borel\sigma$域，则称$g(x)$是一元$Borel$可测函数
  - $\forall B_1\in\mathscr{B}_1$，有$\{\omega:g(\xi(\omega))\in B_1\}=\{\omega:\xi(\omega)\in g^{-1}(B_1)\}\in\mathscr{F}$
  - 离散卷积公式：若ξ与η是相互独立的随机变量，都取非负整数值，其概率分布分别为$\{a_k\}$及$\{b_k\}$，
    则随机变量$\zeta=\xi+\eta$的概率分布为$\{\zeta=r\}=\{\xi=0,\eta=r\}+\{\xi=1,\eta=r-1\}+\cdots+\{\xi=r,\eta=0\}$
    利用独立性得到$c_r=P\{\zeta=r\}=a_0b_r+a_1b_{r-1}+\cdots+a_rb_0,r=1,2\cdots$
- 单个随机变量的函数的分布率
  - ξ的密度函数为$p(x),\eta=g(\xi)$的密度函数是$q(y)$）
    - $\eta=a\xi+b,a\neq0$，求$q(y)$
    - $\eta=\sin\xi$，求$q(y)$
    - $\xi\sim U[-\frac{\pi}{2},\frac{\pi}{2}],\eta=\tan\xi$，求$q(y)$
  - (1)若$g(x)$严格单调，反函数$g^{-1}(y)$连续可导，则$q(y)=p[g^{-1}(y)]|[g^{-1}(y)]'|$
  - (2)若$g(x)$分段严格单调，反函数分别为$h_1(y),h_2(y),\cdots$
    则$q(y)=p[h_1(y)]|h_1'(y)|+p[h_2(y)]|h_2'(y)|+\cdots$

  注意证明
- 均匀分布的特殊地位
- 随机变量的存在定理：若$F(x)\begin{cases}左连续\\单调不减\\F(-\infty)=0,F(+\infty)=1\end{cases}$，则存在一个概率空间$(\Omega,\mathscr{F},P)$，及其上的随机变量$\xi(\omega)$，使$\xi(\omega)$的分布函数正好是$F(x)$

  证明很绕强调过
- 随机向量的函数的分布律
  - 和的分布$\eta=\xi_1+\xi_2$
    - $F_\eta(y)=P\{\eta<y\}=\int_{-\infty}^{+\infty}\int_{-\infty}^{y-x_1}p(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2$
    - 卷积公式

      当ξ_2,ξ_2独立时，$q(y)=\begin{cases}\int_{-\infty}^{+\infty}p_1(u)p_2(y-u)\mathrm{d}u\\ \int_{-\infty}^{+\infty}p_1(y-u)p_2(u)\mathrm{d}u\end{cases}$
  - 商的分布$\eta=\dfrac{\xi_1}{\xi_2}$
    - $F_\eta(x)=P\left\{\dfrac{\xi_1}{\xi_2}<x\right\}=\int_0^{\infty}\left[\int_{-\infty}^{zx}p(y,z)\mathrm{d}y\right]\mathrm{d}z+\int_{-\infty}^{0}\left[\int_{zx}^{\infty}p(y,z)\mathrm{d}y\right]\mathrm{d}z$
    - $q(x)=\int_{-\infty}^{+\infty}|z|p(zx,z)\mathrm{d}z$
  - 关于顺序统计量的若干分布
    - $\xi_1,\cdots,\xi_n$【独立同分布】，$\xi_1^*,\cdots,\xi_n^*$相当于把$\xi_1,\cdots,\xi_n$按照大小重新排列$\xi_1^*\leqslant\xi_2^*\leqslant\cdots\leqslant\xi_n^*$
    - 极大值ξ_n^*的分布函数
      $P\{\xi_n^*<x\}=P\{\max(\xi_1,\xi_2,\cdots,\xi_n)<x\}$
      $=P\{\xi_1<x,\cdots,\xi_n<x\}$
      $=P\{\xi_1<x\}\cdots P\{\xi_n<x\}$
      $=[F(x)]^n$
    - 极小值ξ_1^*的分布函数
      由于$P\{\xi_1^*\geqslant x\}=P\{\min(\xi_1,\xi_2,\cdots,\xi_n)\geqslant x\}$
      $=P\{\xi_1\geqslant x,\cdots,\xi_n\geqslant x\}$
      $=P\{\xi_1\geqslant x\}\cdots P\{\xi_n\geqslant x\}$
      $=[1-F(x)]^n$
      则$P\{\xi_1^*<x\}=1-[1-F(x)]^n$
    - (ξ_1^*,ξ_n^*)的联合分布
      - $G(x,y)=P\{\xi_1^*<x,\xi_n^*<y\}$
        若$x\geqslant y$，则$G(x,y)=P\{\xi_1^*<x,\xi_n^*<y\}=P\{\xi_n^*<y\}=[F(y)]^n$
        若$x<y$，则$G(x,y)=P\{\xi_1^*<x,\xi_n^*<y\}$
        $=P\{\xi_n^*<y\}-P\{\xi_1^*\geqslant x,\xi_n^*<y\}$
        $=[F(y)]^n-[F(y)-F(x)]^n$
      - $q(x,y)=\begin{cases}0, & x\geqslant y\\ n(n-1)[F(y)-F(x)]^{n-2}p(x)p(y), & x<y\end{cases}$
  - 随机变量的变换