# 随机过程期末速通

# 1. 预备知识

## 1.1 概率空间

[**定义1.1.1**] 设$\Omega$是一个样本空间, $\mathcal{F}$是$\Omega$的某些子集组成的集合族, 即所有可测的事件组成的集合. 若满足如下三个条件, 则称$\mathcal{F}$为$\Omega$上的一个$\sigma$**代数**或**域流**, 称$(\Omega,\mathcal{F})$为**可测空间**:

①$\Omega\in\mathcal{F}$.

②[**对补封闭**] 若$A\in\mathcal{F}$, 则$\overline{A}=\Omega\backslash A\in\mathcal{F}$.

③[**对可列并封闭**] 若$A_n\in\mathcal{F}\ \ (n=1,2,\cdots)$, 则$\displaystyle\bigcup_{n=1}^{+\infty}A_n\in\mathcal{F}$.

[**定义1.1.2**] 设$(\Omega,\mathcal{F})$是可测空间, $P(\cdot)$是定义在$\mathcal{F}$上的实值函数, 是测量事件发生的可能性大小的测度. 若满足如下三个条件, 则称$P$是$(\Omega,\mathcal{F})$上的**概率**, 称$(\Omega,\mathcal{F},P)$为**概率空间**, 称$\mathcal{F}$中的元素为**事件**, 对$A\in\mathcal{F}$, 称$P(A)$为事件$A$的**概率**:

①$P(\Omega)=1$.

②对$\forall A\in\mathcal{F}$, 有$0\leq P(A)\leq 1$.

③对一列两两互斥的事件$A_1,A_2,\cdots$, 有$\displaystyle P\left(\bigcup_{i=1}^{+\infty}A_i\right)=\sum_{i=1}^{+\infty}P(A_i)$.

[**定理1.1.1**] 设$(\Omega,\mathcal{F},P)$为概率空间.

(1)[**加法公式**] 若$A,B\in\mathcal{F}$, 有$P(A\bigcup B)+P(A\bigcap B)=P(A)+P(B)$.

(2)[**减法公式**] 若$A,B\in\mathcal{F}$, 且$A\subset B$, 则$P(B-A)=P(B)-P(A)$.

(3)[**单调性**] 若$A,B\in\mathcal{F}$, 且$A\subset B$, 则$P(A)\leq P(B)$.

(4)[**次$\sigma$可加性**] 若$A_n\in\mathcal{F}\ \ (n\geq 1)$, 则$\displaystyle P\left(\bigcup_{n\geq 1}A_n\right)\leq\sum_{n\geq 1}P(A_n)$.

[**注**] (4)中不要求$A_1,A_2,\cdots$互斥.

---

## 1.2 随机变量与分布函数

[**定义1.2.1**] 设$(\Omega,\mathcal{F},P)$是完备的概率空间, $X$是定义在$\Omega$上取值于实数集$\mathbb{R}$的函数. 若对$\forall x\in\mathbb{R}$, 有$\{\omega:X(\omega)\leq x\}=\{\omega\mid X(\omega)\in(-\infty,x]\}\in\mathcal{F}$, 且函数值$\in(-\infty,x]$的原象构成的集合可测, 则称$X$是$\mathcal{F}$上的**随机变量**, 称函数$F(x)=P\{\omega:X(\omega)\leq x\}=P\{\omega:X(\omega)\in(-\infty,x]\}\ \ (-\infty<x<+\infty)$为随机变量$X$的**分布函数**.

[**注1**] $X(\cdot)$是事件(样本点)到实数的映射, $F(\cdot)$是实数到实数的函数.

[**注2**] 注意区间$(-\infty,x]$的右边是闭区间. 利用$F(x)=P\{\omega:X(\omega)\leq x\}$和区间的并、交, 可求得Borel集中其他区间的概率和.

[**注3**] $F(x)$表示比$x$小的概率的累加.

[**注4**] 常用的随机变量分为离散型随机变量和连续型随机变量, 此外还有连续和离散混合的混合型随机变量.

(1)离散型随机变量$X$的取值是有限个或可列个, 其概率分布可用分布列$p_k=P\{X=x_k\}\ \ (k=1,2,\cdots)$描述, 分布函数$F(x)=\sum\limits_{x_k\leq x}p_k$.

(2)连续型随机变量$X$的概率分布用概率密度函数$f(x)$描述, 其分布函数$F(x)=\displaystyle\int_{-\infty}^x f(u)\mathrm{d}u$.

[**定理1.2.1**] 下列四个命题等价, 即一个可测, 其它都可测:

(1)$X$是随机变量.

(2)对$\forall x\in\mathbb{R}$, 都有$\{\omega:X(\omega)\geq x\}\in\mathcal{F}$.

(3)对$\forall x\in\mathbb{R}$, 都有$\{\omega:X(\omega)>x\}\in\mathcal{F}$.

(4)对$\forall x\in\mathbb{R}$, 都有$\{\omega:X(\omega)<x\}\in\mathcal{F}$.

[**定义1.2.2**] 映射$\overrightarrow{X}:\Omega\rightarrow\mathbb{R}^n$表示为$\overrightarrow{X}=(X_1,\cdots,X_n)$. 若对$\forall k\in[1,n]$, $X_k$都是随机变量, 则称$\overrightarrow{X}$为**随机向量**.

[**注1**] 随机向量$\overrightarrow{X}=(X_1,\cdots,X_n)$的联合分布函数 $F(x_1,\cdots,x_n)=P\{X_1\leq x_1,\cdots,X_n\leq x_n\}\ \ (n\geq 1;x_1,\cdots,x_n\in\mathbb{R})$.

[**注2**] 随机向量是静态的, 区别于动态的随机过程.

[**定义1.2.3**] 常见分布.

(1)若随机变量$X$只取常数$c$, 即$P\{X=c\}=1$, 则$X$并不随机, 但将其视为随机变量的退化情况更方便, 称其为**退化分布**或**单点分布**.

(2)在一次试验中, 设事件$A$出现的概率为$p\ \ (0<p<1)$, 不出现的概率为$(1-p)$. 设$X$为事件$A$出现的次数, 则$X$可能的取值为$0,1$, 对应的概率$P\{X=k\}=p^k(1-p)^{1-k}\ \ (k=0,1)$, 称其为**Bernoulli分布**或**两点分布**.

(3)设$n$重Bernoulli试验中, 事件$A$在每次试验中出现的概率都为$p\ \ (0<p<1)$. 设$X$为事件$A$出现的次数, 则$X$可能的取值为$0,1,\cdots,n$, 对应的概率$P\{X=k\}=C_n^k p^k(1-p)^{n-k}\ \ (k=0,1,\cdots,n)$, 称其为以$n$和$p$为参数的**二项分布**, 记作$X\sim B(n,p)$. Bernouliii分布可看作$n=1$时的二项分布, 对应于一次试验的情形.

(4)若随机变量$X$可取一切非负整数值, 且对应的概率$P\{X=k\}=\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}\ \ (k=0,1,\cdots;\lambda>0)$, 则称$X$服从参数为$\lambda$的**Poisson分布**, 记作$X\sim P(\lambda)$.

(5)在Bernoulli试验序列中, 设事件$A$在每次试验中出现的概率都为$p\ \ (0<p<1)$. 设$X$为事件$A$首次出现的试验次数, 则$X$可能的取值为$1,2,\cdots$, 对应的概率$P\{X=k\}=p(1-p)^{k-1}\ \ (k=1,2,\cdots)$, 称其为**几何分布**.

(6)若概率密度函数$f(x)=\begin{cases}\dfrac{1}{b-a},a\leq x<b\\ 0,otherwise\end{cases}\ \ (a<b)$, 则称其为区间$[a,b]$上的**连续均匀分布**, 简称**均匀分布**.

(7)若概率密度函数$f(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\cdot\exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)\ \ (x\in\mathbb{R})$, 则称其为以$\mu$和$\sigma$为参数的**正态分布**或**Gauss分布**, 记作$X\sim N(\mu,\sigma^2)$.

(8)若概率密度函数$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x},x\geq 0\\ 0,x<0\end{cases}$, 则称其为以 $\lambda\ \ (\lambda\neq 0)$ 为参数的**指数分布**.

## 1.3 数字特征、矩母函数与特征函数

**[定义1.3.1]** 对离散型随机变量$X$, 定义其期望$E(X)=\sum\limits_k x_k\cdot P\{X=x_k\}$. 对连续性随机变量$X$, 定义其**期望**$E(X)=\int_{-\infty}^{+\infty}x\cdot f(x)\mathrm{d}x$. 两者可统一为$E(X)=\int_{-\infty}^{+\infty}x\mathrm{d}F(x)=\int_{\omega\in\Omega}X(\omega)P(\mathrm{d}\omega)=\int_{\omega\in\Omega}X(\omega)\mathrm{d}P(\omega)$, 其中$P(\mathrm{d}\omega)=P\{\omega:\omega\in\mathrm{d}\omega\}$.

**[注]** $\int_{-\infty}^{+\infty}x\cdot f(x)\mathrm{d}x$是对$x\in\mathbb{R}$分割, 是Riemann积分, 而$\int_{\omega\in\Omega}X(\omega)P(\mathrm{d}\omega)$是对$\Omega$分割, 是Lebesgue积分.

**[定义1.3.2]** 对随机变量$X$, 若$E\{[X-E(X)]^2\}$存在, 则称其为$X$的**方差**, 记作$Var(X)$, 即$Var(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2\mathrm{d}F(x)$.

**[定义1.3.3]** 对随机变量$X$和$Y$, 若$E\{[X-E(X)][Y-E(Y)]\}$存在, 则称其为$X$与$Y$的**协方差**, 记作$Cov(X,Y)$, 即$Cov(X,Y)=E(XY)-E(X)E(Y)$.

**[定义1.3.4]** 设随机变量$X$的分布函数为$F(x)$, 称函数$\phi(a)=E\left(\mathrm{e}^{aX}\right)=\int_\Omega\mathrm{e}^{aX(\omega)}P(\mathrm{d}\omega)=\int_{-\infty}^{+\infty}\mathrm{e}^{ax}\mathrm{d}F(x)$为$X$的**矩母函数**, 其中$a$为参数.

**[注1]** 矩母函数是关于随机变量函数的期望.

**[注2]** 矩母函数未必存在. 上述积分存在时, 矩母函数存在. 矩母函数存在时唯一决定分布.

**[定理1.3.1]** 设随机变量$X$的矩母函数为$\phi(a)$, 则 $\varphi(a)$ 的 $n$ 阶导在 $a=0$ 处的值 $\varphi^{(n)}(0)=E(X^n)$.

**[注]** 设$g(X(\omega))=\mathrm{e}^{aX(\omega)}$, 则$E\left(\mathrm{e}^{aX(\omega)}\right)=E[g(X(\omega))]=\int_{-\infty}^{+\infty}g(x)\mathrm{d}F(x)=\int_{\omega\in\Omega}g(X(\omega))P(\mathrm{d}\omega)$.

可以证明对$\phi(a)$求导时, 求导运算和求期望(积分)运算可换序, 由复合函数求导法则可得$X$的各阶矩:

(1)$\phi'(a)=E\left[\mathrm{e}^{aX(\omega)}\cdot X(\omega)\right]$, 则$X$的一阶矩为$\phi'(a)\Big|_{a=0}=E[X(\omega)]$.

(2)$\phi''(a)=E\left\{\left[\mathrm{e}^{aX(\omega)}\cdot[X(\omega)]^2\right]\right\}$, 则$X$的二阶矩为$\phi''(a)\Big|_{a=0}=E\{[X(\omega)]^2\}$.

**[定义1.3.5]** 设随机变量$X$的分布函数为$F(x)$, 称函数$\psi(a)=E\left(\mathrm{e}^{iaX}\right)=\int_\Omega\mathrm{e}^{iaX(\omega)}\mathrm{d}P(\omega)=\int_{-\infty}^{+\infty}\mathrm{e}^{iax}\mathrm{d}F(x)$为$X$的**特征函数**.

**[注1]** 特征函数是一个实变复值函数.

**[注2]** 不同于矩母函数, 随机变量的特征函数必存在.

[**定理1.3.2**] 设随机变量$X$的特征函数为$\psi(u)$.

(1)作线性变换$Y=aX+b$, 则随机变量$Y$的特征函数$\psi_Y(u)=\mathrm{e}^{ibu}\psi_X(au)$.

(2)两相互独立的随机变量之和的特征函数等于它们各自的特征函数之积.

(3)若$X$存在$n$阶矩, 则其特征函数可微分$n$次, 且对$\forall k\leq n$, 有$\psi^{(k)}(0)=i^kE(X^k)$.

[**证**]

(1)$\psi_Y(u)=E\left(\mathrm{e}^{iuY}\right)=E\left(\mathrm{e}^{iu(aX+b)}\right)=E\left(\mathrm{e}^{iuaX}\cdot\mathrm{e}^{iub}\right)=\mathrm{e}^{iub}\cdot E\left(\mathrm{e}^{iuaX}\right)$.

(2)设随机变量$X_1$与$X_2$相互独立,

则$\psi_{X_1+X_2}(u)=E\left(\mathrm{e}^{iu(X_1+X_2)}\right)=E\left(\mathrm{e}^{iuX_1}\right)\cdot E\left(\mathrm{e}^{iuX_2}\right)=\psi_{X_1}(u)\cdot\psi_{X_2}(u)$.

[**定理1.3.3**] [**唯一性定理**] 分布函数由其特征函数唯一确定.

[**推论**] 分布函数与特征函数相互唯一确定.

---

---

## 1.4 独立性

参考概率论.

---

---

## 1.5 条件期望

[**定义1.5.1**]

(1)设$X$和$Y$都是离散型随机变量. 对所有 $s.t.\ P\{Y=y\}>0$的$y$, 在$Y=y$的前提下$X$的**条件概率** $P\{X=x\mid Y=y\}=\dfrac{P\{X=x,Y=y\}}{P\{Y=y\}}$, **条件分布函数**$F(x\mid y)=P\{X\leq x\mid Y=y\}$, **条件期望** $E(X\mid Y=y)=\displaystyle\int x\mathrm{d}F(x\mid y)=\sum_x x\cdot P\{X=x\mid Y=y\}$.

(2)设连续型随机变量$X$和$Y$有联合概率密度函数$f(x,y)$. 对所有 $s.t.\ f_Y(y)\geq 0$的$y$, 在$Y=y$的前提下$X$的**条件概率密度**$f(x\mid y)=\dfrac{f(x,y)}{f_Y(y)}$, **条件分布函数**$F(x\mid y)=P\{X\leq x\mid Y=y\}=\displaystyle\int_{-\infty}^x f(u\mid y)\mathrm{d}u$, **条件期望** $E(X\mid Y=y)=\displaystyle\int x\mathrm{d}F(x\mid y)=\int x\cdot f(x\mid y)\mathrm{d}x$.

(3)离散型、连续型随机变量的条件期望可统一为$E(X\mid Y=y)=\displaystyle\int_{-\infty}^{+\infty}x\cdot\mathrm{d}F(x\mid Y=y)$.

[**注**] $E(X\mid Y)$是关于$Y$的函数, 当$Y=y$时取值$E(X\mid Y=y)$.

[**定理1.5.1**] 设$X$和$Y$都是随机变量. 若随机变量$E(X \mid Y)$的期望存在, 则$E[E(X \mid Y)] = E(X)$.

[**证**] 下面只证明$X$为离散型随机变量的情况.

$$
\begin{aligned}
E[E(X \mid Y)] &= \sum_j \sum_i x_j \cdot P\{X = x_j \mid Y = y_i\} \cdot P\{Y = y_i\} \\
&= \sum_i \sum_j x_j \cdot \frac{P\{X = x_j, Y = y_i\}}{P\{Y = y_i\}} \cdot P\{Y = y_i\} \\
&= \sum_j x_j \cdot \sum_i P\{X = x_j \mid Y = y_i\} = \sum_j x_j \cdot P\{X = x_j\} = E(X).
\end{aligned}
$$

[**注**] 直观上, $E[E(X \mid Y)]$表示先对区域求平均, 再对各区域的结果求平均, 结果与直接对整体求平均相同.

[**定理1.5.2**] [**全期望公式**] 设$X$和$Y$都是随机变量. 若随机变量$E(X \mid Y)$的期望存在, 则

$$
E\{E[X(\omega) \mid Y(\omega)]\} = \begin{cases} \text{离散型: } \displaystyle\sum_i E[X(\omega) \mid Y = y_i] \cdot P\{Y = y_i\} \\ \text{连续型: } \displaystyle\int_{-\infty}^{+\infty} E[X(\omega) \mid Y = y] \cdot f_Y(y) \mathrm{d}y \end{cases} = \int_{-\infty}^{+\infty} E(X \mid Y = y) \mathrm{d}F_Y(y).
$$

[**注**] 连续型表达式中的$E[X(\omega) \mid Y = y]$中的$Y = y$实际上为$Y \in \mathrm{d}y$.

[**例1.5.1**] [**随机个随机变量之和**] 设$X_1, X_2, \cdots$是一列独立同分布的随机变量, $N$是一个非负整值随机变量(与时间无关), 且与序列$X_1, X_2, \cdots$独立. 求随机变量$Y = \sum\limits_{i=1}^{N} X_i$的均值和方差.

[**解**] 因$X_i$和$N$都随机, 考虑在固定$N$的前提下求$Y$的矩母函数.

$$
\phi_Y(a) = E\left(\mathrm{e}^{aY}\right) = E\left[\exp\left(a \sum_{i=1}^{N} X_i\right)\right] = E\left\{E\left[\exp\left(a \sum_{i=1}^{N} X_i\right) \middle| N\right]\right\}.
$$
*注意最后不是$N = n$.

考察$N = n$时的情况.

$$
\begin{aligned}
&E\left[\exp\left(a \sum_{i=1}^{N} X_i\right) \middle| N = n\right] = E\left[\exp\left(a \sum_{i=1}^{n} X_i\right) \middle| N = n\right] = E\left(\mathrm{e}^{aX_1} \times \cdots \times \mathrm{e}^{aX_n}\right) \\
&\xlongequal{X_1, \cdots, X_n\text{独立同分布}} \left[E\left(\mathrm{e}^{aX_1}\right)\right]^n = [\phi_X(a)]^n.
\end{aligned}
$$

则$E\left[\exp\left(a \sum\limits_{i=1}^{N} X_i\right) \middle| N\right] = [\phi_X(a)]^N$, 进而$\varPhi_Y(a) = E\{[\phi_X(a)]^N\}$.

可以证明求导和求期望可交换, 则$\phi_Y'(a) = E\{N \cdot [\phi_X(a)]^{N-1} \cdot \phi_X'(a)\}$.

令$a = 0$, 则$E(Y) = E[N \cdot E(X)] = E(N) \cdot E(X)$.

同理$\phi_Y''(a) = E\{N(N-1) \cdot [\phi_X(a)]^{N-2} \cdot [\phi_X'(a)]^2 + N \cdot [\phi_X(a)]^{N-1} \cdot \phi_X''(a)\}$.

令$a = 0$, 则$E(Y^2) = E\{N(N-1) \cdot [E(X)]^2 + N \cdot E(X^2)\} = E(N) \cdot Var(X) + E(N^2) \cdot [E(X)]^2$.

故$Var(Y^2) = E(Y^2) - [E(Y)]^2 = E(N) \cdot Var(X) + [E(X)]^2 \cdot Var(N)$.

[**注**] $E[N \cdot E(X)] = E(N) \cdot E(X)$称为**Wald公式**.

**[例1.5.2]** 某矿工被困在有三个门的矿井中, 第一个门走$2\ \mathrm{h}$可达出口, 第二个门走$3\ \mathrm{h}$会回到原处, 第三个门走$5\ \mathrm{h}$会回到原处. 设该矿工等可能地在三个门中任选一个, 他到达出口的时间为$X$. 求$X$的均值和矩母函数.

**[解]**

(1)设该矿工第一次选的门为$Y$, 则事件$\{Y=i\}\ \ (1\leq i\leq 3)$表示他第一次选$i$号门.

因$P\{Y=1\}=P\{Y=2\}=P\{Y=3\}=\dfrac{1}{3}$,

则$E(X)=E[E(X\mid Y)]=\dfrac{E(X\mid Y=1)+E(X\mid Y=2)+E(X\mid Y=3)}{3}$

$=\dfrac{2+[3+E(X)]+[5+E(X)]}{3}$, 解得: $E(X)=10$.

(2)$E\left(\mathrm{e}^{tX}\right)=E\left[E\left(\mathrm{e}^{tX}\mid Y\right)\right]=\dfrac{E\left(\mathrm{e}^{tX}\mid Y=1\right)+E\left(\mathrm{e}^{tX}\mid Y=2\right)+E\left(\mathrm{e}^{tX}\mid Y=3\right)}{3}$.

显然$E\left(\mathrm{e}^{tX}\mid Y=1\right)=\mathrm{e}^{2t}$.

$Y=2$时, 设该矿工第一次选择2号门走$3\ \mathrm{h}$回到原处后还需$X'\ \mathrm{h}$到达出口, 则$X=3+X'$.

因$X\overset{d}{=}X'$, 则$E\left(\mathrm{e}^{tX}\mid Y=2\right)=E\left(\mathrm{e}^{t(3+X')}\right)=\mathrm{e}^{3t}\cdot E\left(\mathrm{e}^{tX}\right)$.

同理$E\left(\mathrm{e}^{tX}\mid Y=3\right)=E\left(\mathrm{e}^{t(5+X')}\right)=\mathrm{e}^{5t}\cdot E\left(\mathrm{e}^{tX}\right)$.

故$E\left(\mathrm{e}^{tX}\right)=\dfrac{\mathrm{e}^{2t}+\mathrm{e}^{3t}\cdot E\left(\mathrm{e}^{tX}\right)+\mathrm{e}^{5t}\cdot E\left(\mathrm{e}^{tX}\right)}{3}=\dfrac{\mathrm{e}^{2t}}{3-\mathrm{e}^{3t}-\mathrm{e}^{5t}}$.

**[定理1.5.3]** 设$X$和$Y$都是随机变量, 子$\sigma$代数$\mathcal{G}\subset\mathcal{F}$, 则:

(1)**[重期望公式]** $E[E(X\mid\mathcal{G})]=E(X)$.

(2)若$X$是$\mathcal{G}$可测的($\mathcal{G}$已知时, 可认为$X$是常数), 则$E(X\mid G)=X$, $\mathrm{a.s.}$.

(3)若$\mathcal{G}=\{\varnothing,\Omega\}$, 则$E(X\mid\mathcal{G})=E(X)$, a.s., 即空集和样本空间不提供额外的信息.

(4)若$X$和随机变量$XY$的期望都存在, 且$Y$是$\mathcal{G}$可测的, 则$E(XY\mid\mathcal{G})=Y\cdot E(X\mid\mathcal{G})$, a.s..

(5)若$X$与$\mathcal{G}$相互独立(即$\sigma(X)$与$\mathcal{G}$相互独立), 则$E(X\mid\mathcal{G})=E(X)$, a.s..

(6)**[条件期望的平滑性]** 若$\mathcal{G}_1,\mathcal{G}_2$是两个子$\sigma$代数, 且$\mathcal{G}_1\subset\mathcal{G}_2\subset\mathcal{F}$, 则$E[E(X\mid\mathcal{G}_2)\mid\mathcal{G}_1]=E(X\mid\mathcal{G}_1)$, a.s..

**[注1]** $\mathrm{a.s.}$为almost surely, 即几乎必然成立, 表示成立的概率为$1$, 不成立的概率为$0$(不成立的点集是零测集).

**[注2]** (4)是(2)的加强, (3)是(5)的特例.