销售一个零件的平均利润为

$$\begin{aligned}E(T)&=-P\{T=-1\}+20P\{T=20\}+(-5)P\{T=-5\}\\&=-\Phi(10-\mu)+20[\Phi(12-\mu)-\Phi(10-\mu)]-5[1-\Phi(12-\mu)]\\&=25\Phi(12-\mu)-21\Phi(10-\mu)-5.\end{aligned}$$

因为

$$\begin{aligned}\frac{\mathrm{d}E(T)}{\mathrm{d}\mu}&=-25\varphi(12-\mu)+21\varphi(10-\mu)\\&=-25\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(12-\mu)^2}{2}}+21\times\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(10-\mu)^2}{2}},\end{aligned}$$

令$\frac{\mathrm{d}E(T)}{\mathrm{d}t}=0$，得

$$\mathrm{e}^{-(22-2\mu)}=\frac{21}{25}.$$

解得

$$\mu=11+\frac{1}{2}\ln\frac{21}{25}\approx10.9.$$

容易验证 $\mu=11+\frac{1}{2}\ln\frac{21}{25}$是$E(T)$的最大值点，即当 $\mu\approx10.9$ (mm)时，销售一个零件的平均利润最大，此最大值为

$$\begin{aligned}E(T)\Big|_{\mu=10.9}&=25\Phi(12-10.9)-21\Phi(10-10.9)-5\\&=25\times0.8643-21\times(1-0.8159)-5\approx12.74\ (\text{元}).\end{aligned}$$

例 4.32 假设一部机器在一天内发生故障的概率为 0.2，机器发生故障时全天停止工作．若一周 5 个工作日里无故障，可获利润 10 万元；发生一次故障仍可获利润 5 万元；发生二次故障所获利润 0 元；发生三次或三次以上故障就要亏损 2 万元．求一周内期望利润是多少？

［提示］ 本题考查离散型随机变量的数学期望．首先要正确写出相应的概率分布．

［解］ 以 X 表示一周 5 天内机器发生故障天数，则 X 服从参数为(5,0.2)的二项分布：

$$P\{X=k\}=\mathrm{C}_5^k0.2^k0.8^{5-k},k=0,1,2,3,4,5.$$

$$P\{X=0\}=0.8^5=0.328,$$

$$P\{X=1\}=\mathrm{C}_5^1(0.2)(0.8)^4=0.410,$$

$$P\{X=2\}=\mathrm{C}_5^2(0.2)^2(0.8)^3=0.025,$$

$$P\{X\geqslant3\}=1-P\{X=0\}-P\{X=1\}-P\{X=2\}=0.057.$$

以 Y 表示所获利润，则

$$Y=g(X)=\begin{cases}10, & \text{若 } X=0,\\ 5, & \text{若 } X=1,\\ 0, & \text{若 } X=2,\\ -2, & \text{若 } X\geqslant3.\end{cases}$$

$$E(Y)=10\times0.328+5\times0.410+0\times0.205-2\times0.057=5.216\ (\text{万元}).$$

五、大数定律和中心极限定理

·考试内容与要求·

考试内容

切比雪夫(Chebyshev)不等式　切比雪夫大数定律　伯努利(Bernoulli)大数定律　辛钦(Khinchine)大数定律　棣莫弗－拉普拉斯(De Moivre-Laplace)定理　列维－林德伯格(Levy-Lindberg)定理

考试要求

1. 了解切比雪夫不等式.

2. 了解切比雪夫大数定律、伯努利大数定律和辛钦大数定律(独立同分布随机变量序列的大数定律).

3. 了解棣莫弗－拉普拉斯定理(二项分布以正态分布为极限分布)和列维－林德伯格定理(独立同分布随机变量序列的中心极限定理).

·考试内容解析·

(一) 切比雪夫不等式

设随机变量 X 的方差 $D(X)$ 存在，则对于任意给定的正数 ε，总有

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2},$$

或

$$P\{|X-E(X)|<\varepsilon\}\geqslant 1-\frac{D(X)}{\varepsilon^2}.$$

(二) 大数定律

1. 依概率收敛

设 $X_1,X_2,\cdots,X_n,\cdots$ 是一个随机变量序列，a 是一个常数. 如果对于任意给定的正数 ε、有

$$\lim_{n\to\infty}P\{|X_n-a|<\varepsilon\}=1,$$

则称随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$ 依概率收敛于 a，记作

$$X_n\xrightarrow{P}a.$$

如果两个随机变量序列 $X_n\xrightarrow{P}a$，$Y_n\xrightarrow{P}b$. 函数 $g(a,b)$ 在点 (a,b) 连续. 则

$$g(X_n,Y_n)\xrightarrow{P}g(a,b).$$

2. 切比雪夫大数定律

设 $X_1,X_2,\cdots,X_n,\cdots$ 是由两两不相关(或两两独立)的随机变量所构成的序列，分别具有数学期望

$$E(X_1),E(X_2),\cdots,E(X_n),\cdots$$

和方差

$$D(X_1),D(X_2),\cdots,D(X_n),\cdots,$$

并且方差有公共上界，即存在正数 M，使得

$$D(X_n)\leqslant M,\quad n=1,2,\cdots,$$

则对于任意给定的正数 ε，总有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\right|<\varepsilon\right\}=1.$$

3. 独立同分布的切比雪夫大数定律

设随机变量 $X_1,X_2,\cdots,X_n,\cdots$ 相互独立(即对任意正整数 $k>1$,随机变量 $X_1,X_2,\cdots,X_k$ 是相互独立的)，服从相同的分布，具有数学期望 $E(X_n)=\mu$ 和方差 $D(X_n)=\sigma^2$ $(n=1,2,\cdots)$. 则对于任意给定的正数 ε，总有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|<\varepsilon\right\}=1.$$

即随机变量序列 $\overline{X}_n=\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{P}\mu$.

4. 伯努利大数定律

设在每次试验中事件 A 发生的概率 $P(A)=p$，在 n 次独立重复试验中，事件 A 发生的频率为 $f_n(A)$，则对于任意正数 ε，总有

$$\lim_{n\to\infty} P\{|f_n(A)-p|<\varepsilon\}=1.$$

伯努利大数定律表明：在相同条件下进行大量(n 次)独立重复试验时，随机事件 A 发生的频率 $f_n(A)$ 稳定在事件 A 的概率 $P(A)$ 附近，因此在试验次数 n 充分大时，可以用频率 $f_n(A)$ 作为概率 $P(A)$ 的近似值。如果概率 $P(A)$ 很小，则事件 A 发生的频率也很小，由此可知在一次试验中事件 A 几乎不可能发生。这也是人们在长期实践中总结出来的经验，称为实际推断原理或小概率事件的实际不可能原理，有着广泛的应用。如果随机事件的概率接近于 1，则在一次试验中这一事件几乎一定发生。

5. 辛钦大数定律

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，具有数学期望 $E(X_n)=\mu\ (n=1,2,\cdots)$，则对于任意给定的正数 ε，总有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k-\mu\right|<\varepsilon\right\}=1.$$

(三) 中心极限定理

1. 棣莫弗－拉普拉斯中心极限定理

设随机变量 X_n 服从参数为 n 和 p 的二项分布，即 $X_n\sim B(n,p)(0<p<1, n=1,2,\cdots)$，则对于任意实数 x，有

$$\lim_{n\to\infty} P\left\{\frac{X_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\Phi(x).$$

在定理的条件下，对于任意整数 $k\ (0\leqslant k\leqslant n)$，当 n 充分大时，有

$$C_n^k p^k(1-p)^{n-k}\approx\frac{1}{\sqrt{2\pi np(1-p)}}e^{-\frac{(k-np)^2}{2np(1-p)}}=\frac{1}{\sqrt{np(1-p)}}\varphi\left(\frac{k-np}{\sqrt{np(1-p)}}\right).$$

对于任意整数 k_1 和 $k_2\ (0\leqslant k_1<k_2\leqslant n)$，当 n 充分大时，有

$$\sum_{k=k_1}^{k_2} C_n^k p^k(1-p)^{n-k}\approx\int_a^b\varphi(x)\,dx=\Phi(b)-\Phi(a),$$

对于任意有限区间 $[c,d]$，当 n 充分大时，有

$$P\left\{c<\frac{X_n-np}{\sqrt{np(1-p)}}\leqslant d\right\}\approx\int_c^d\varphi(x)\,dx=\Phi(d)-\Phi(c).$$

其中 $a=\dfrac{k_1-np}{\sqrt{np(1-p)}}$，$b=\dfrac{k_2-np}{\sqrt{np(1-p)}}$，$\varphi(x)$ 和 $\Phi(x)$ 分别为服从标准正态分布的随机变量的概率密度和分布函数。第一个结果称为局部极限定理，后两个结果称为积分极限定理。

2. 列维－林德伯格中心极限定理(独立同分布的中心极限定理)

设随机变量 $X_1, X_2, \cdots, X_n, \cdots$ 相互独立，服从同一分布，具有数学期望和方差

$$E(X_n)=\mu,\quad D(X_n)=\sigma^2>0,\quad n=1,2,\cdots,$$

则对于任意实数 x，有

$$\lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} X_k-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\Phi(x).$$

在定理的条件下，当 n 充分大时，$\sum_{k=1}^{n} X_k$ 的标准化随机变量

$$\frac{\sum_{k=1}^{n} X_k-n\mu}{\sqrt{n}\sigma}=\frac{\frac{1}{n}\sum_{k=1}^{n} X_k-\mu}{\sigma/\sqrt{n}}$$

近似服从标准正态分布 $N(0,1)$。

・例题详解・

例5.1　设随机变量 X 的方差为 2，则根据切比雪夫不等式有估计 $P\{|X-E(X)|\geqslant 2\}\leqslant$______.

［答案］ $\frac{1}{2}$.

［提示］ 本题考查切比雪夫不等式的应用.

［解］ 切比雪夫不等式为：对于任意 $\varepsilon>0$,

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant D(X)/\varepsilon^2.$$

令 $\varepsilon=2$, $D(X)=2$, 则 $P\{|X-E(X)|\geqslant 2\}\leqslant\frac{D(X)}{2^2}=\frac{1}{2}$.

［典型错误］ 对于切比雪夫不等式或极限定理中有关的内容，可能有不少考生在复习时未予重视，从而对此看来十分简单的填空，一片茫然，束手无策.

例5.2 设随机变量 X 和 Y 的数学期望分别为 -2 和 2，方差分别为 1 和 4，而相关系数为 -0.5，则根据切比雪夫不等式 $P\{|X+Y|\geqslant 6\}\leqslant$______.

［答案］ $\frac{1}{12}$.

［提示］ 本题涉及的知识点包括两个随机变量和的数学期望、方差和相关系数.

［解］ 由切比雪夫不等式，对于任意 $\varepsilon>0$,

$$P\{|X-E(X)|\geqslant\varepsilon\}\leqslant\frac{D(X)}{\varepsilon^2}.$$

由于 $E(X+Y)=E(X)+E(Y)=-2+2=0$,

$$D(X+Y)=D(X)+D(Y)+2\rho_{XY}D(X)D(Y)$$
$$=1+4-2\times0.5\times\sqrt{1\times4}=3,$$

令 $\varepsilon=1$.

则
$$P\{|X+Y|\geqslant 6\}\leqslant\frac{D(X+Y)}{36}=\frac{3}{36}=\frac{1}{12}.$$

例5.3 设总体 X 服从参数为 2 的指数分布，$x_1,x_2,\cdots,x_n$ 为来自总体 X 的简单随机样本，则当 $n\to\infty$ 时 $Y_n=\frac{1}{n}\sum\limits_{i=1}^{n}X_i^2$ 依概率收敛于______.

［答案］ $\frac{1}{2}$.

［提示］ 本题考查考生对独立同分布大数定理概念的理解及运用的能力，并要求考生熟练掌握，指数分布的数学期望，方差及其间的运算关系.

［解］ 由于 $x_1,x_2,\cdots,x_n$ 是来自总体 X 的简单随机样本，因而 $x_1,x_2,\cdots,x_n$ 相互独立，并可以推出 $x_1^2,x_2^2,\cdots,x_n^2$ 也相互独立并且同分布. 又因 X 服从参数为 2 的指数分布，所以

$$E(X_i)=\frac{1}{2},\ D(X_i)=\frac{1}{4}\quad(i=1,2,\cdots,n),$$

则
$$E(X_i^2)=D(X_i)+[E(X_i)]^2=\frac{1}{4}+\left(\frac{1}{2}\right)^2=\frac{1}{2}\quad(i=1,2,\cdots,n).$$

由独立同分布大数定律知

$$\frac{1}{n}\sum_{i=1}^{n}X_i^2\text{ 依概率收敛于}\frac{1}{2}.$$

例5.4 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，$S_n=X_1+X_2+\cdots+X_n$，则根据列维－林德伯格(Levy-Lindberg)中心极限定理，当 n 充分大时，S_n 近似服从正态分布，只要 $X_1,X_2,\cdots,X_n$

(A) 有相同的数学期望.　　(B) 有相同的方差.

(C) 服从同一指数分布.　　(D) 服从同一离散型分布.

［答案］ (C).

［提示］ 本题考查列维－林德伯格定理所应满足的条件.

［解］ 列维－林德伯格中心极限定理：设 $X_1,X_2,\cdots,X_n$ 是独立同分布的随机变量，$E(X_i)=\mu$，$D(X_i)=\sigma^2(i=1,2,\cdots,n)$，则 S_n 近似服从正态分布 $N(n\mu,n\sigma^2)$. 对照该定理，(A)、(B)、(C)、(D)4 个

选项中只有(C)符合条件，因为$X_1,X_2,\cdots,X_n$服从同一指数分布时，$E(X_i)$和$D(X_i)$均存在.

(A)不成立，因为X_i有相同的数学期望，并不能保证X_i方差存在；同样，X_i有相同的方差，但数学期望不一定存在，如X服从柯西分布，X_i服从同一离散分布也不能保证数学期望和方差均存在.

例5.5 设$X_1,X_2,\cdots,X_n,\cdots$为独立同分布的随机变量序列，且均服从参数为$\lambda\ (\lambda>1)$的指数分布，记$\Phi(x)$为标准正态分布函数，则

(A) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\lambda\sqrt{n}}\leqslant x\right\}=\Phi(x)$.　　(B) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\sqrt{n\lambda}}\leqslant x\right\}=\Phi(x)$.

(C) $\lim\limits_{n\to\infty}P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n}X_i-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$.　　(D) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-\lambda}{\sqrt{n\lambda}}\leqslant x\right\}=\Phi(x)$.

[提示] 本题考查中心极限定理的应用.

[答案] (C).

[解] 由题设知，$\mu=\dfrac{1}{\lambda}$，$\sigma^2=\dfrac{1}{\lambda^2}$，$\sigma=\dfrac{1}{\lambda}$，故由独立同分布中心极限定理知应选(C).

例5.6 某保险公司多年的统计资料表明，在索赔户中被盗索赔户占20%，以X表示在随意抽查的100个索赔户中因被盗向保险公司索赔的户数

(1) 写出X的概率分布；

(2) 利用棣莫佛－拉普拉斯定理，求被盗索赔户不少于14户且不多于30户的概率的近似值。

附表：设$\Phi(x)$是标准正态分布函数

x	0	0.5	1.0	1.5	2.0	2.5	3.0
$\Phi(x)$	0.500	0.692	0.841	0.933	0.977	0.994	0.999

[提示] 本题考查二项分布及棣莫佛－拉普拉斯定理的应用，属于简单计算题.

[解] (1) X服从二项分布，参数$n=100$，$p=0.2$

$$P\{X=k\}=C_{100}^{k}0.2^{k}0.8^{100-k},\quad k=0,1,\cdots,100.$$

(2) $E(X)=np=20$，$D(X)=np(1-p)=16$.

根据棣莫佛－拉普拉斯定理

$$\begin{aligned}P\{14\leqslant X\leqslant 30\}&=P\left\{\frac{14-20}{4}\leqslant\frac{X-20}{4}\leqslant\frac{30-20}{4}\right\}\\&=P\left\{-1.5\leqslant\frac{X-20}{4}\leqslant 2.5\right\}\approx\Phi(2.5)-\Phi(-1.5)\\&=\Phi(2.5)-[1-\Phi(1.5)]=\Phi(2.5)+\Phi(1.5)-1\\&=0.944+0.933-1=0.927.\end{aligned}$$

例5.7 假设$X_1,X_2,\cdots,X_n$是来自总体X的简单随机样本；已知$E(X^k)=\alpha_k\ (k=1,2,3,4)$. 证明当$n$充分大时，随机变量$Z_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2$近似服从正态分布，并指出其分布参数.

[提示] 根据题设条件，应用中心极限定理——独立同分布的中心极限定理，即可得证.

[证明] 依题意$X_1,X_2,\cdots,X_n$独立同分布，可知$X_1^2,X_2^2,\cdots,X_n^2$也独立同分布且有

$$E(X_i^2)=\alpha_2,\quad D(X_i^2)=E(X_i^4)-[E(X_i^2)]^2=\alpha_4-\alpha_2^2,$$

由中心极限定理

$$V_n=\frac{\sum\limits_{i=1}^{n}X_i^2-n\alpha_2}{\sqrt{n(\alpha_4-\alpha_2^2)}}=\frac{\frac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}=\frac{Z_n-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}$$

的极限分布是标准正态分布. 所以当n充分大时V_n近似服从标准正态分布，从而$Z_n=\sqrt{\dfrac{\alpha_4-\alpha_2^2}{n}}\cdot V_n+\alpha_2$近

似服从参数为 $\mu=a_2$，$\sigma^2=\dfrac{a_4-a_2^2}{n}$ 的正态分布.

例5.8 一生产线生产的产品成箱包装，每箱的重量是随机的. 假设每箱平均重 50 千克，标准差为 5 千克. 若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多可以装多少箱，才能保障不超载的概率大于 0.977. ($\Phi(2)=0.977$. 其中 $\Phi(x)$是标准正态分布函数.)

[提示] 中心极限定理是针对 n 个独立同分布的随机变量和而言的，根据题意构造这样的一个独立随机变量和，最后求出 n.

[解] 设 $X_i\ (i=1,2,\cdots,n)$是装运的第 i 箱的重量(单位:千克)，n 是所求箱数. 由条件可以把 X_1，$X_2,\cdots,X_n$ 视为独立同分布随机变量，则 n 箱的总重量

$$T_n = X_1+X_2+\cdots+X_n$$

是独立同分布随机变量之和.

由条件知 $E(X_i)=50$，$\sqrt{D(X_i)}=5$；$E(T_n)=50n$，$\sqrt{D(T_n)}=5\sqrt{n}$(单位:千克).

根据列维－林德伯格中心极限定理，T_n 近似服从正态分布 $N(50n,25n)$.

箱数 n 决定于条件

$$P\{T_n\leqslant 5\,000\}=P\left\{\frac{T_n-50n}{5\sqrt{n}}\leqslant\frac{5\,000-50n}{5\sqrt{n}}\right\}$$

$$\approx\Phi\left(\frac{1\,000-10n}{\sqrt{n}}\right)>0.977=\Phi(2).$$

由此可见

$$\frac{1\,000-10n}{\sqrt{n}}>2,$$

从而 $n<98.019\,9$，即最多可以装 98 箱.

六、数理统计的基本概念

·考试内容与要求·

考试内容

总体　个体　简单随机样本　统计量　样本均值　样本方差和样本矩　χ^2 分布　t 分布　F 分布　分位数　正态总体的常用抽样分布

考试要求

1. 理解总体、简单随机样本、统计量、样本均值、样本方差及样本矩的概念. 其中样本方差定义为:

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2.$$

2. 了解 χ^2 分布、t 分布和 F 分布的概念及性质，了解分位数的概念并会查表计算.
3. 了解正态总体的常用抽样分布.

·考试内容解析·

(一) 总体和样本

1. 总体

在数理统计中所研究对象的某项数量指标 X 取值的全体称为总体，X 是一个随机变量. X 的分布函数和数字特征分别称为总体的分布函数和数字特征. 总体中的每个元素称为个体. 每个个体是一个实数. 总体中个体的数量称为总体的容量. 容量为有限的总体称为有限总体，容量为无限的总体称为无限总体.

2. 简单随机样本

设总体 X 的分布函数为 $F(x)$，如果随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，并且都具有分布函数 $F(x)$，则称 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的简单随机样本，简称为样本，n 称为样本容量. 它们的观测值 $x_1,x_2,\cdots,x_n$