第 5 章 概率统计基础

1. 设总体 X 的概率密度为 $f(x)=\begin{cases}\dfrac{x}{2}, & 0<x<2,\\ 0, & 其他.\end{cases}$，$X_1,X_2,X_3,X_4$ 为来自总体的容量为 4 的简单随机样本，求 X_1,X_2,X_3,X_4 的联合概率密度.

【分析】 利用简单随机样本联合概率密度的定义.

解 X_1,X_2,X_3,X_4 的联合概率密度为

$$f(x_1,x_2,x_3,x_4)=\prod_{i=1}^{4}f(x_i)=\begin{cases}\left(\dfrac{1}{2}\right)^4 x_1x_2x_3x_4, & 0<x_1,x_2,x_3,x_4<2,\\ 0, & 其他.\end{cases}$$

2. 设总体 X 在区间 $[a,b]$ 上服从均匀分布，$X_1,X_2,\cdots,X_n$ 为来自总体的简单随机样本，$x_1,x_2,\cdots,x_n$ 为样本值，求 $X_1,X_2,\cdots,X_n$ 的联合概率密度.

【分析】利用均匀分布的定义和简单随机样本联合概率密度的定义.

解 总体的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a\le x\le b,\\ 0, & 其他.\end{cases}$$

则 $X_1,X_2,\cdots,X_n$ 的联合概率密度为

$$f(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i)=\begin{cases}\left(\dfrac{1}{b-a}\right)^n, & a\le x_1,x_2,\cdots,x_n\le b,\\ 0, & 其他.\end{cases}$$

3. 设一批灯泡的寿命 X (单位：万小时)服从参数为 6 的指数分布,从该批灯泡中采用简单随机抽样抽取容量为 10 的样本 $X_1,X_2,\cdots,X_{10}$. 求 $X_1,X_2,\cdots,X_{10}$ 的联合概率密度.

【分析】利用指数分布的定义和简单随机样本联合概率密度的定义.

解 总体的概率密度为

$$f(x)=\begin{cases}\frac{1}{6}\mathrm{e}^{-\frac{1}{6}x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

则 $X_1,X_2,\cdots,X_{10}$ 的联合概率密度为

$$f(x_1,x_2,\cdots,x_{10})=f(x_1)f(x_2)\cdots f(x_{10})=\begin{cases}\left(\frac{1}{6}\right)^{10}\mathrm{e}^{-\frac{1}{6}\sum_{i=1}^{10}x_i}, & x_1,x_2,\cdots,x_{10}>0,\\ 0, & \text{其他}.\end{cases}$$

4. 在总体 $X\sim N(10,4)$ 中抽取一容量为 10 的简单随机样本 X_1,X_2,X_3,X_4,X_5，求 $P\{\max(X_1,X_2,\cdots,X_{10})>12\}$.

【分析】利用独立随机变量的性质和概率计算公式.

解 $P\{\max(X_1,X_2,\cdots,X_{10})>12\}$

$$=1-P\{\max(X_1,X_2,\cdots,X_{10})\le 12\}$$

$$=1-P\{X_1\le 12,\cdots,X_{10}\le 12\}$$

$$=1-[P(X\le 12)]^{10}$$

$$=1-[\Phi(1)]^{10}$$

5. 某公司 40 名职工月基本工资（单位：元）如下：

2210	2500	2480	3100	3700	2100	2900	2240	2350	2860
3210	2350	2450	2390	2700	1180	2200	1580	1890	1620
2960	2720	2700	2380	3590	1920	2550	2490	2370	2420
2880	2450	2430	3270	2470	2420	2530	2570	2600	2620

绘制月工资的频率直方图.

【分析】参考例题，按步骤画图即可.

解 (1)决定组数、组距和组限

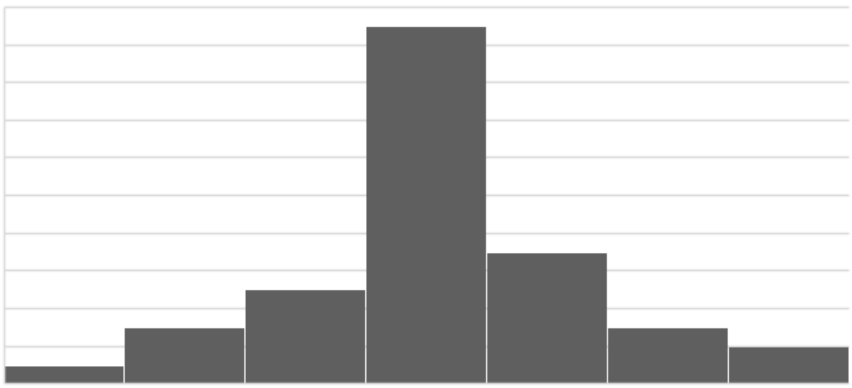
数据里面的极小值为 1180，极大值为 3700。即所有数据落在区间[1180，3700]。取区间 $I=[1179.5,3700.5]$。将数据分为 7 组，对应把区间等分为 7 个小区间，于是小区间长度为 $(3700.5-1179.5)/7\approx 360.143$，为了方便计算，将区间 I 进行调整为[1162.5，3717.5]组距为 $\Delta=(3717.5-1162.5)/7=365$。

(2)列频数分布表为：

数出落在每个小区间内的数据的频数 f_i，算出频率 $f_i/n(n=40, i=1,2,\cdots,7)$，列出如下频数分布表：

组限	频数 f_i	频率 f_i/n	累积频率
$1162.5\sim1527.5$	1	0.025	0.025
$1527.5\sim1892.5$	3	0.075	0.1
$1892.5\sim2257.5$	5	0.125	0.225
$2257.5\sim2622.5$	19	0.475	0.7
$2622.5\sim2987.5$	7	0.175	0.875
$2987.5\sim3352.5$	3	0.075	0.95
$3352.5\sim3717.5$	2	0.05	1

(3)画直方图（图形需要请编辑们添加一下字，与教材参考答案相同）



6. 某厂从一批荧光灯中抽出 10 个，测其寿命的数据（单位：千小时）如下：

95.5，　18.1，　13.1，　26.5，　31.7，　33.8，　8.7，　15.0，　48.8，　48.3

求该批荧光灯寿命的经验分布函数 $F_n(x)$.

【分析】利用经验分布函数的定义.

解 将数据由小到大排列得：

8.7，13.1，15.0，18.1，26.5，31.7，33.8，48.8，49.3，95.5，

则经验分布函数为：

$$F_n(x)=\begin{cases}0, & x<8.7,\\ 0.1, & 8.7\le x<13.1,\\ 0.2, & 13.1\le x<15.0,\\ 0.3, & 15.0\le x<18.1,\\ 0.4, & 18.1\le x<26.5,\\ 0.5, & 26.5\le x<31.7,\\ 0.6, & 31.7\le x<33.8,\\ 0.7, & 33.8\le x<48.8,\\ 0.8, & 48.8\le x<49.3,\\ 0.9, & 49.3\le x<95.5,\\ 1, & x\ge 95.5.\end{cases}$$

7. 设总体 X 服从正态分布 $N(\mu,\sigma^2)$，其中 μ 已知，σ^2 未知，X_1,X_2,X_3 是取自总体 X 的一个样本，下列表达式哪些是统计量，哪些不是统计量？为什么？

(A) $\frac{1}{5}(X_1+X_2+X_3)$　　　　(B) $X_1+X_2+2\mu$

(C) $\max(X_1,X_2,X_3)$　　　　(D) $\frac{1}{\sigma^2}(X_1+X_2+X_3)$

【分析】利用统计量的定义.

解 (A) (B) (C)是统计量；(D)不是统计量，因为含有未知参数 σ^2.

8. 为了研究某种零件的定额加工工时，随机观察了 12 人次的加工工时，测得如下数据(单位：分钟):

$$9.8,7.8,8.2,10.5,7.5,8.8,10.0,9.4,8.5,9.5,8.4,9.8,$$

试求样本均值、样本方差、样本标准差.

【分析】利用样本均值、样本方差、样本标准差的定义.

解 $\bar{x}=\frac{1}{12}\sum_{k=1}^{12}x_k\approx 9.0167,\quad s^2=\frac{1}{12-1}\sum_{k=1}^{12}(x_k-\bar{x})^2\approx 0.9015,\quad s=\sqrt{s^2}\approx 0.9495$。

9. 给定一组样本 $X_1,X_2,\cdots,X_9$，其观测值为 $x_1,x_2,\cdots,x_9$，且有 $\sum_{i=1}^{9}x_i=45,\sum_{i=1}^{9}x_i^2=285,$ 则样本方差 S^2 的观测值为是多少？

【分析】利用样本方差的定义.

解 $$s^2=\frac{\sum_{i=1}^{9}(x_i-\bar{x})^2}{9-1}=\frac{\sum_{i=1}^{9}x_i^2-9\times\bar{x}^2}{9-1}=\frac{285-9\times 25}{8}=7.5$$

10. 设总体 X 服从(0-1)分布，抽取样本 $X_1,X_2,\cdots,X_n$，求样本平均值 $\bar{X}$ 的数学期望

$E(\bar{X})$ 及方差 $D(\bar{X})$.

【分析】利用方差和标准差的性质.

解 可以证明 $n\bar{X} \sim B(n,p)$，所以

$$E(\bar{X}) = \frac{1}{n}E(n\bar{X}) = \frac{1}{n}\times np = p,$$
$$D(\bar{X}) = \frac{1}{n^2}D(n\bar{X}) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

11. 设 $X \sim N(1,1), Y \sim N(1,4)$，$X$ 与 Y 相互独立，证明 $(X-1)^2 + (Y-1)^2/4 \sim \chi^2(2)$.

【分析】利用正态分布的标准化和 χ^2 分布的定义.

解 因为 $X \sim N(1,1), Y \sim N(1,4)$，

所以 $X-1 \sim N(0,1), (Y-1)/2 \sim N(0,1)$，且相互独立，

根据定义 $(X-1)^2 + (Y-1)^2/4 \sim \chi^2(2)$.

12. 设随机变量 X 和 Y 相互独立，且都服从正态分布 $N(0,3^2)$,设 $X_1, X_2, \cdots, X_9$ 和 $Y_1, Y_2, \cdots, Y_9$ 分别是来自两个总体的简单随机样本,证明统计量 $U = \dfrac{\sum_{i=1}^{9} X_i}{\sqrt{\sum_{i=1}^{9} Y_i^2}}$ 服从分布 $t(9)$.

【分析】利用相正态分布随机变量的线性组合的性质、χ^2 分布的定义和 t 分布的定义.

解 $\sum_{i=1}^{9} X_i \sim N(0,9^2) \Rightarrow \sum_{i=1}^{9} X_i \Big/ 9 \sim N(0,1)$，$\sum_{i=1}^{9} Y_i^2 \Big/ 9 \sim \chi^2(9)$

由 t 分布的定义有 $U = \dfrac{\sum_{i=1}^{9} X_i}{\sqrt{\sum_{i=1}^{9} Y_i^2}} = \dfrac{\sum_{i=1}^{9} X_i \Big/ 9}{\sqrt{\sum_{i=1}^{9} Y_i^2 \Big/ 81}} \sim t(9)$

13. 从正态总体 $N(\mu,\sigma^2)$ 中抽取容量为 16 的样本，已知 $\sigma^2 = 25$，试求样本均值 $\bar{X}$ 与总体均值 μ 之差的绝对值小于 2 的概率.

【分析】利用方差已知的正态总体期望的分布.

解 在 σ^2 已知时，统计量

$$u = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

因此，

$$P\left\{\left|\bar{X}-\mu\right|<2\right\}=P\left\{\frac{\left|\bar{X}-\mu\right|}{\sigma/\sqrt{n}}<\frac{2}{\sigma/\sqrt{n}}\right\}=P\left\{\frac{\left|\bar{X}-\mu\right|}{5/4}<\frac{2\times 4}{5}\right\}$$

$$=P\left\{|u|<1.6\right\}=\Phi(1.6)-\Phi(-1.6)=2\Phi(1.6)-1=2\times 0.9452-1=0.8904.$$

14. 在总体 $N(\mu,\sigma^2)$ 中随机抽取一容量为16的简单随机样本(其中 μ,σ 均未知)，求 $P(S^2/\sigma^2\le 2.041)$ 及 $E(S^2),D(S^2)$.

【分析】利用方差未知的正态总体方差的分布以及期望和方差的性质.

解 由于 $X_1,X_2,...,X_{16}$ 是总体 $N(\mu,\sigma^2)$ 的简单随机样本，故 $15S^2/\sigma^2\sim\chi^2(15)$，故

$$P(S^2/\sigma^2\le 2.041)=1-P(15S^2/\sigma^2>30.615)\approx 1-0.01=0.99,$$

$$E(S^2)=\frac{\sigma^2}{15}E(15S^2/\sigma^2)=\frac{\sigma^2}{15}\times 15=\sigma^2,$$

$$D(S^2)=(\frac{\sigma^2}{15})^2D(15S^2/\sigma^2)=\frac{\sigma^4}{225}\times 30=\frac{2\sigma^4}{15}.$$