第一章 统计量与抽样分布

1.1 基本概念

1.2 充分统计量与完备统计量

1.3 抽样分布

1.4 次序统计量及其分布

1.1 基本概念

一、总体和样本

二、统计量和样本矩

三、经验分布函数

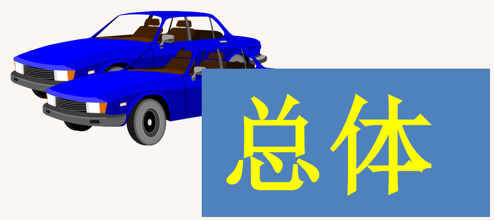
一、总体与样本

1. 总体（Population）

一个统计问题总有它明确的研究对象.

研究对象的全体元素组成的集合称为总体(母体)，总体中每个成员称为个体.

研究某批灯泡的质量



…

考察国产 轿车的质量

然而在统计研究中，人们关心总体仅仅是关心其每个个体的一项(或几项)数量指标和该数量指标在总体中的分布情况．这时，每个个体具有的数量指标的全体就是总体.

该批灯泡寿命的全体就是总体

所有国产轿车每公里耗油量的全体就是总体

而每个个体具有的数量指标的全体往往形成一定的概率分布.

考察某大学一年级学生的年龄

某大学一年级全体学生的年龄构成问题的总体

设该大学一年级学生的年龄分布如下表

年龄	18	19	20	21	22
比例	0.5	0.3	0.1	0.07	0.03

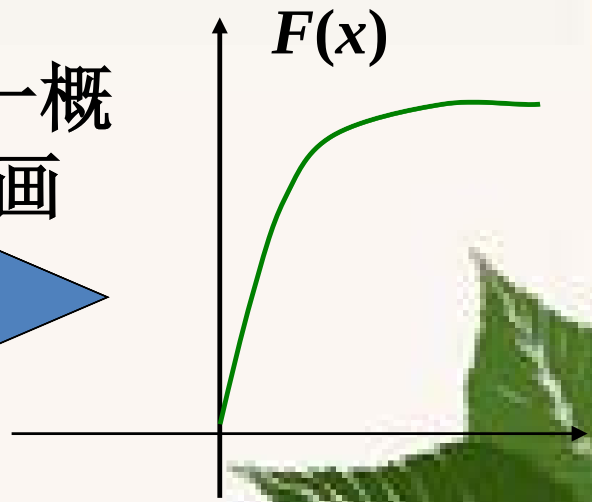
显然，上述分布是一个概率分布，记作X.

既然*r.v.* *X*的概率分布和总体的各个值的分布一样，即*X*的概率分布可以反映总体中各个值的分布情况。

自然，我们就用随机变量*X*或者随机变量*X*的分布来表示所考察的总体.

又如:研究某批灯泡的寿命时，所有灯泡的寿命值形成一个总体，所有灯泡的寿命值也会形成一定的概率分布，此概率分布用某个随机变量X或用其分布函数$F(x)$描述.于是，这个寿命总体就可用随机变量X或用其$F(x)$ 刻画，记为X.

寿命X可用一概率分布来刻画



结论 **总体可以用一个随机变量及其分布来描述.**

因此，常用随机变量的记号或用其分布函数表示总体. 比如说总体X或总体$F(x)$.

比如说正态总体X，表示总体X的分布是正态分布。

2. 样本（Sample）

从总体X中，随机抽取n个个体：

$$X_1, X_2, \cdots, X_n$$

称为总体X的一个**样本**，记为

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

n称为**样本容量**.

注

每一个个体X_i $(i=1,2,\cdots,n)$都是一个随机变量。
样本$(X_1, X_2, \cdots, X_n)$是一个n维随机变量.

考察某大学一年级学生的年龄

某大学一年级全体学生的年龄构成问题的总体

设该大学一年级学生的年龄分布如下表(即总体X的分布如下）

年龄	18	19	20	21	22
比例	0.5	0.3	0.1	0.07	0.03

若从该大学一年级学生中任意抽查一个学生的年龄，记作X_1.

则X_1的概率分布是：

X_1	18	19	20	21	22
p	0.5	0.3	0.1	0.07	0.03

可见，X_1是一个随机变量。同理，X_i $(i=2,3,\ldots,n)$都是随机变量。

3. 样本值

每次抽取 $X_1, X_2, \cdots, X_n$ 所得到的n个确定的具体数值，记为（$x_1, x_2, \cdots, x_n$）称为样本 $(X_1, X_2, \cdots, X_n)$ 的一个**样本值**(观察值).

4. 简单随机样本

若来自总体X的样本$(X_1, X_2, \cdots, X_n)$具有下列

两个**特征**：

(1) 代表性： $X_1, X_2, \ldots, X_n$中每一个与所考察的总体有相同的分布.

(2) 独立性： $X_1, X_2, \ldots, X_n$是相互独立的随机变量.

则称$(X_1, X_2, \cdots, X_n)$是来自总体X，容量为n的简单随机样本.

获得简单随机样本的抽样方法称为简单随机抽样.

5. 总体和样本的严格数学定义:

总体 一个随机变量X或其相应的分布函数$F(x)$称为一个总体. $F(x)$

样本 设X是具有分布函数$F(x)$的随机变量,若$X_1, X_2, \cdots, X_n$是具有同一分布函数$F(x)$、相互独立的随机变量,则称$X_1, X_2, \cdots, X_n$为来自总体X的容量为n的简单随机样本,简称样本.

6. 样本的分布

定理1.1 设$(X_1,X_2,\cdots,X_n)$为来自总体X的样本.

(1) 若总体$X\sim F(x)$, 则样本

$$(X_1 X_2,\cdots,X_n)\sim\prod_{i=1}^{n}F(x_i)$$

(2)若总体$X\sim p(x)$,则样本

$$(X_1,X_2,\cdots,X_n)\sim\prod_{i=1}^{n}p(x_i)$$

(3)若总体$X\sim P(X=x_i^*)=p(x_i^*)(i=1,2,\cdots)$

则样本$(X_1,X_2,\cdots,X_n)\sim\prod_{i=1}^{n}p(x_i)$.其中每一个$x_1,x_2,\cdots,x_n$都取$x_1^*,x_2^*,\cdots$中的任意值.

熟记3种不同情形下样本的联合概率分布的具体形式，是解决问题的关键 ！！

例1 设总体 X 服从参数为 $\lambda\,(\lambda>0)$ 的指数分布, $(X_1,X_2,\cdots,X_n)$ 是来自总体的样本, 求样本 $(X_1,X_2,\cdots,X_n)$ 的概率密度.

证 总体 X 的概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$

所以 $(X_1,X_2,\cdots,X_n)$ 的概率密度为

$$f(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i)=\begin{cases}\prod_{i=1}^{n}\lambda e^{-\lambda x_i}, & x_i>0\\ 0, & \text{其它}\end{cases}$$

$$=\begin{cases}\lambda^n e^{-\lambda\sum_{i=1}^{n}x_i}, & x_i>0\ (i=1,2,\cdots,n)\\ 0, & \text{其它}\end{cases}$$

例2 设总体X服从P (λ), $(X_1, X_2, \ldots, X_n)$是来自于X的样本，求此样本的联合分布律.

证 总体 X的分布律为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,\cdots$$

因为 $X_1, X_2, \cdots, X_n$ 独立同分布,

所以, $(X_1, X_2, \cdots, X_n)$ 的联合分布律为

$$P(X_1=k_1,\cdots,X_n=k_n)=\prod_{i=1}^{n}P(X_i=k_i)$$

注意离散型样本的联合分布律的表达形式。

$$P(X_1 = k_1, \cdots, X_n = k_n) = \prod_{i=1}^{n} P(X_i = k_i)$$

$$= \prod_{i=1}^{n} \frac{\lambda^{k_i}}{k_i\,!} e^{-\lambda} = \frac{\lambda^{k_1 + \cdots + k_n}}{k_1\,! k_2\,! \cdots k_n\,!} e^{-n\lambda}$$

$$k_i = 0, 1, 2, \cdots, i = 1, 2, \cdots, n$$

此即样本的联合分布律

二、统计量与样本矩

由样本推断总体情况,需要对样本值进行"加工",这就需要构造一些样本的函数,它把样本中所含的信息集中起来.

1. 统计量的定义

定义1.1　设$(X_1, X_2, \cdots, X_n)^T$是来自总体$X$一个样本, $f(X_1, X_2, \cdots, X_n)$是$X_1, X_2, \cdots, X_n$的函数, 若f中不含任何关于总体X的未知参数, 则称$f(X_1, X_2, \cdots, X_n)$是一个统计量.

设$(x_1, x_2, \cdots, x_n)^T$是样本$(X_1, X_2, \cdots, X_n)^T$的样本值,

则称$f(x_1, x_2, \cdots, x_n)$是$f(X_1, X_2, \cdots, X_n)$的观察值.

用于估计分布中参数的统计量，称为**估计量.**

注 1° 统计量$f(X_1, X_2, \cdots, X_n)$是随机变量;

2° 统计量用于统计推断，故不应含任何关于总体X的未知参数.

例3 设 X_1, X_2, X_3 是来自总体 $N(\mu,\sigma^2)$ 的一个样本, 其中 μ 为已知, σ^2 为未知, 判断下列各式哪些是统计量, 哪些不是?

$$T_1 = X_1, \qquad T_4 = \max(X_1, X_2, X_3),$$

$$T_2 = X_1 + X_2 e^{X_3}, \qquad T_5 = X_1 + X_2 - 2\mu,$$

$$T_3 = \frac{1}{3}(X_1 + X_2 + X_3),$$

是

$$T_6 = \frac{1}{\sigma^2}(X_1^2 + X_2^2 + X_3^2).$$

不是

2. 三种常用统计量

(1) 样本矩

设 $X_1, X_2, \cdots, X_n$ 是来自总体的一个样本，

$x_1, x_2, \cdots, x_n$ 是这一样本的观察值.

1) **样本均值** $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i;$

2) **样本方差**

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

3) 样本标准差

$$S_n=\sqrt{S_n^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2}\,;$$

4) 修正样本方差

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right).$$

S_n^2与S_n^{*2}的关系 $S_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{n-1}{n}S_n^{*2}.$

注 1° 当n较大时，S_n^{*2}与S_n^2差别微小；

2° 当n较小时，S_n^{*2}比S_n^2有更好的统计性质.

5) 样本 k 阶(原点)矩

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \text{ , } k = 1, 2, \cdots;$$ 特例：$A_1 = \overline{X}$

6)样本 k 阶中心矩 特例：$B_2 = S_n^2$

$$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k \text{ , } k = 2, 3, \cdots;$$

样本矩的意义

推断

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \longrightarrow EX=\mu$$

反映总体均值的信息

$$S_n^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 \longrightarrow DX=\sigma^2$$

$$S_n^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2 \longrightarrow DX=\sigma^2$$

反映总体方差的信息

$$S_n=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2} \longrightarrow \sqrt{DX}=\sigma$$

$$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \longrightarrow E(X^k)=\mu_k$$

$$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k \longrightarrow E(X-EX)^k=\beta_k$$

3、样本矩的性质

定理1.2 设总体X具有$2k$阶矩，则来自总体X的样本k阶原点距A_k的数学期望和方差分别为：

$$E(A_k)=\alpha_k \qquad D(A_k)=\frac{\alpha_{2k}-\alpha_k^2}{n}$$

其中$\alpha_k = E(X^k)(k=1,2,\cdots,)$表示总体的$k$阶原点距.

证 $$E(A_k)=E(\frac{1}{n}\sum_{i=1}^{n}X_i^k)=\frac{1}{n}\sum_{i=1}^{n}E(X_i^k)=\frac{1}{n}\sum_{i=1}^{n}E(X^k)=\alpha_k$$

$$D(A_k)=D(\frac{1}{n}\sum_{i=1}^{n}X_i^k)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i^k)=\frac{1}{n^2}\sum_{i=1}^{n}D(X^k)$$

$$=\frac{1}{n}((E(X^{2k})-E^2X^k)=\frac{\alpha_{2k}-\alpha_k^2}{n}$$

推论 设总体X的期望$EX=\mu$,方差$DX=\sigma^2$,k阶矩EX^k也存在,$(X_1,X_2,\cdots,X_n)$为来自总体X的样本,则有

(1) $E(\overline{X})=\mu$

(2) $D(\overline{X})=\dfrac{1}{n}\sigma^2$;

(3) $E(S_n^2)=\dfrac{n-1}{n}\sigma^2$;

(4) $E(S_n^{*2})=\sigma^2$.

这些结论对任何总体都适用，请熟记这些结论！！！！！

证

(1) $E(\overline{X})=\mu$

$$E(\overline{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}\mu=\mu$$

(2) $D(\overline{X})=\frac{1}{n}\sigma^2$

$$D(\overline{X})=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=\frac{1}{n}\sigma^2.$$

（3） $E(S_n{}^2)=\dfrac{n-1}{n}\sigma^2$

$$E(S_n^2)=E[\frac{1}{n}\sum_{i=1}^{n}X_i^2-\overline{X}^2]=\frac{1}{n}\sum_{i=1}^{n}E(X_i^2)-E(\overline{X}^2)$$

$$=\frac{1}{n}\sum_{i=1}^{n}[D(X_i)+(E(X_i))^2]-[D(\overline{X})+(E(\overline{X}))^2]$$

$$=\frac{1}{n}\sum_{i=1}^{n}(\sigma^2+\mu^2)-(\frac{1}{n}\sigma^2+\mu^2)=\frac{n-1}{n}\sigma^2.$$

(4) $E(S_n^{*2})=E(\frac{n}{n-1}S_n^2)=\frac{n}{n-1}E(S_n^2)=\sigma^2$

例4 设总体 $X \sim B(1, p)$,

$(X_1, X_2, \cdots, X_n)$ 是来自于总体的样本，

$\bar{X}$和S_n^2是样本均值与样本方差，试计算：

$E\bar{X}, D\bar{X}$和ES_n^2.

证 由两点分布知 $EX = p, DX = p(1-p)$

利用样本矩的性质得

$$E\bar{X} = EX = p$$

$$D\bar{X} = \frac{1}{n}DX = \frac{p(1-p)}{n}$$

$$ES_n^2 = \frac{n-1}{n}DX = \frac{n-1}{n}p(1-p).$$

样本矩的极限性质

若总体X的k阶矩$E(X^k)=\mu_k$存在，

则当$n\to\infty$时，$A_k \xrightarrow{P} \mu_k, k=1,2,\cdots$

证 因为 $X_1, X_2, \cdots, X_n$ 独立且与 X 同分布，

所以 $X_1^k, X_2^k, \cdots, X_n^k$ 独立且与 X^k 同分布，

故有 $E(X_1^k)=E(X_2^k)=\cdots=E(X_n^k)=\mu_k.$

根据辛钦大数定律，可得

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{P} \mu_k, \quad k=1,2,\cdots.$$

注 此性质是下一章矩估计法的理论根据.

结论 $\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow[n\to\infty]{P} \mu_k$ ， $k = 1, 2, \cdots.$

2个特例

(1) $k = 1$, $A_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} \mu = EX$,

(2) $k = 2$, $A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow[n\to\infty]{P} \mu_2 = EX^2$,

进而 根据依概率收敛的序列的性质知

$$\boxed{g(A_1, A_2, \cdots, A_k) \xrightarrow[n\to\infty]{P} g(\mu_1, \mu_2, \cdots, \mu_k),}$$

其中g是连续函数.

2个结论

(1) $$\underline{A_2 - A_1^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \underline{S_n^2}$$

$$\xrightarrow[n\to\infty]{P} \mu_2 - \mu^2 = EX^2 - (EX)^2 = DX = \sigma^2 .$$

(2) $$S_n^{*2} = \frac{n}{n-1} S_n^2 \xrightarrow[n\to\infty]{P} DX = \sigma^2 .$$

三、经验分布函数

1、次序统计量

设$(X_1, X_2, \cdots, X_n)^T$是从总体$X$中抽取的一个样本，$(x_1, x_2, \cdots, x_n)^T$是其一个观测值，将观测值按由小到大的次序重新排列为

$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$$

当$(X_1, X_2, \cdots, X_n)^T$取值为$(x_1, x_2, \cdots x_n)^T$时，定义$X_{(k)}$取值为$x_{(k)}$ $(k = 1, 2, \cdots n)$，由此得到

$$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})^T$$

称为样本$(X_1, X_2, \cdots, X_n)^T$**的次序统计量.**

对应的$(x_{(1)}, x_{(2)}, \cdots x_{(n)})$称为其观测值.

2、经验分布函数

定义1.3 设 $X_1, X_2, \cdots, X_n$ 是总体 X 的一个样本，

$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 为总体 X 的样本 $(X_1, X_2, \cdots, X_n)$ 的次序统计量.

$(x_{(1)}, x_{(2)}, \cdots x_{(n)})$ 为其观测值，设 x 是任一实数，

称函数

$$F_n(x) = \begin{cases} 0, & x < x_{(1)}, \\ \dfrac{k}{n}, & x_{(k)} \le x < x_{(k+1)}, \\ 1, & x \ge x_{(n)}. \end{cases}$$

为总体 X 的经验分布函数。

例5 设总体X具有一个样本值1, 2, 1.5, 3,

排序 1, 1.5，2, 3

则经验分布函数

$F_3(x)$的观察值为 $$F_3(x)=\begin{cases}0, & x<1,\\ \dfrac{1}{4}, & 1\le x<1.5,\\ \dfrac{2}{4}, & 1.5\le x<2,\\ \dfrac{3}{4}, & 2\le x<3,\\ 1, & x\ge 3.\end{cases}$$

理解:

对于任意实数x,$F_n(x)$可理解为为样本值中不超过x的个数再除以n,亦即

$$F_n(x)=\frac{\upsilon_n(x)}{n}$$

其中$\upsilon_n(x)$表示 $x_1,x_2,\cdots,x_n$ 中不超过 于 x 的个数.

注: $\upsilon_n(x)$ 可看做是n次重复独立试验中,事件$\{X\le x\}$发生的次数.称为经验频数。

$$F_n(x)=\frac{\upsilon_n(x)}{n}$$

为$\{X\le x\}$的频率.

$$P\{\upsilon_n(x)=k\}=C_n^k(P\{X\le x\})^k(1-P\{X\le x\})^{n-k}$$

$$=C_n^kF^k(x)(1-F(x))^{n-k}$$

即 $\upsilon_n(x)\sim B(n,F(x))$

3、经验分布函数的性质

(1)对于给定的一组样本值$(x_1, x_2, \cdots, x_n)$，$F_n(x)$满足分布函数的特征,是一个分布函数.

(2)由于$F_n(x)$是样本的函数,故$F_n(x)$是随机变量.

由 $\upsilon_n(x) \sim B(n, F(x))$ 及$F_n(x) = \dfrac{\upsilon_n(x)}{n}$,

可得 $nF_n(x) \sim B(n, F(x))$,所以

$$E[F_n(x)] = F(x), \quad D[F_n(x)] = \frac{F(x)[1 - F(x)]}{n}$$

(3) $F_n(x)$依概率收敛于$F(x)$. 即

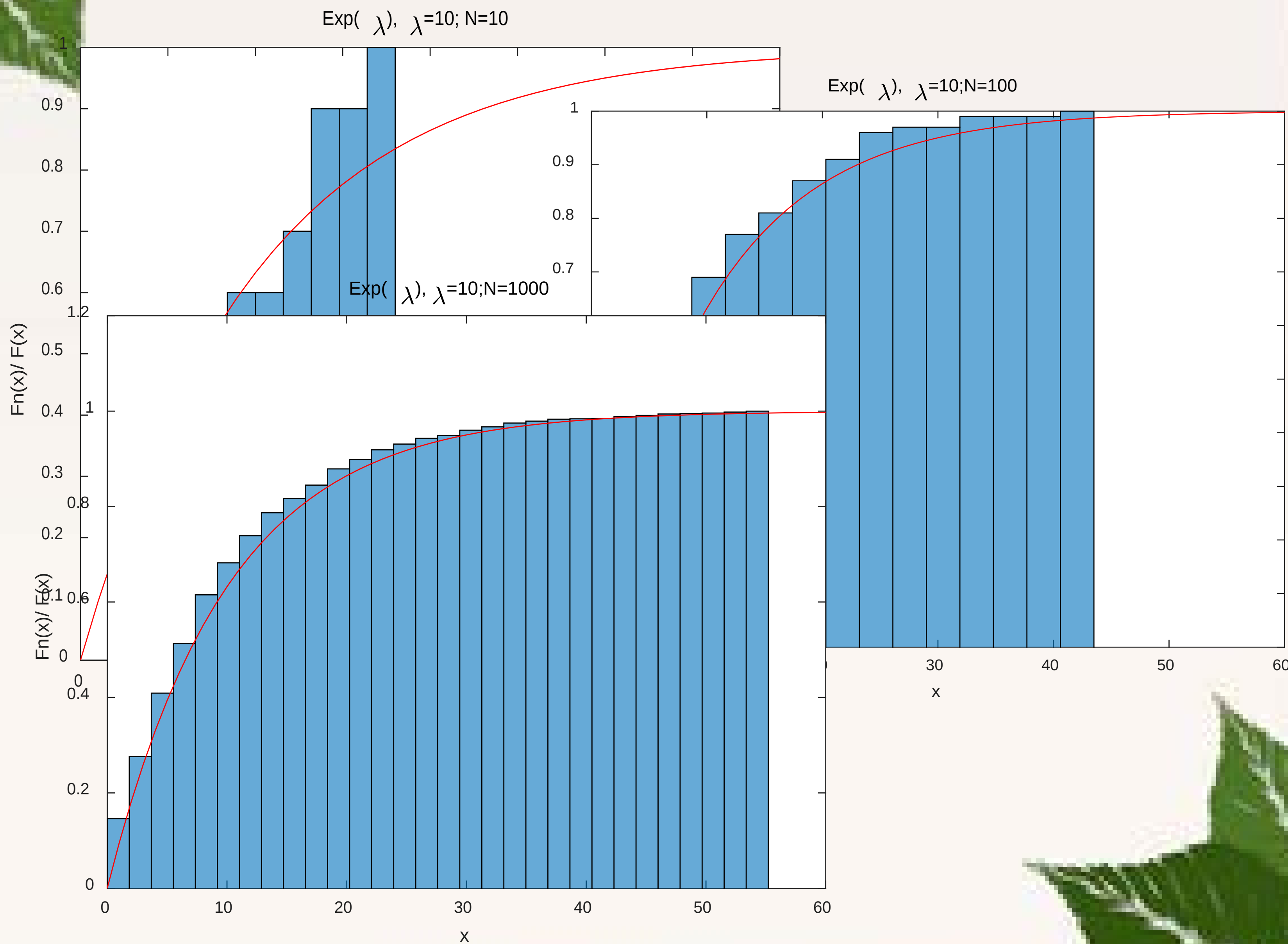
$$\lim_{n\to\infty} P\{|F_n(x)-F(x)|<\varepsilon\}=1 \qquad (\forall\varepsilon>0)$$

（4）格里汶科定理

对于任一实数 x, 当$n\to\infty$时, $F_n(x)$以概率1一致收敛于分布函数 $F(x)$, 即

$$P\left\{\lim_{n\to\infty}\sup_{-\infty<x<+\infty}\left|F_n(x)-F(x)\right|=0\right\}=1.$$

对于$\forall x\in R$,当n充分大时，$F_n(x)$ 与总体分布函数$F(x)$只有微小的差别，从而实际中,当$F(x)$未知时，可通过样本构造一个$F_n(x)$，近似当做$F(x)$来使用。


Exp(λ), λ=10; N=10
Exp(λ), λ=10;N=100
Exp(λ), λ=10;N=1000
Fn(x)/ F(x)
Fn(x)/ F(x)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1.2
0
10
20
30
40
50
60
x
x

填空题 1分

设置

经验分布函数是统计量吗？ [填空1]

正常使用填空题需3.0以上版本雨课堂

作答

Thank You!

辛钦定理

设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立，服从同一分布，且具有数学期望$E(X_k) = \mu\ (k = 1, 2, \cdots)$,

则对于任意正数ε,有$\lim\limits_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1.$

格里汶科资料

Boris Vladimirovich Gnedenko

Born: 1 Jan 1912 in Simbirsk (now Ulyanovskaya), Russia
Died: 27 Dec 1995 in Moscow, Russia