

# 心理统计学第三讲

讲师

司马紫衣



高途学院APP下载



高途学院公众号





# 第六章 概率分布

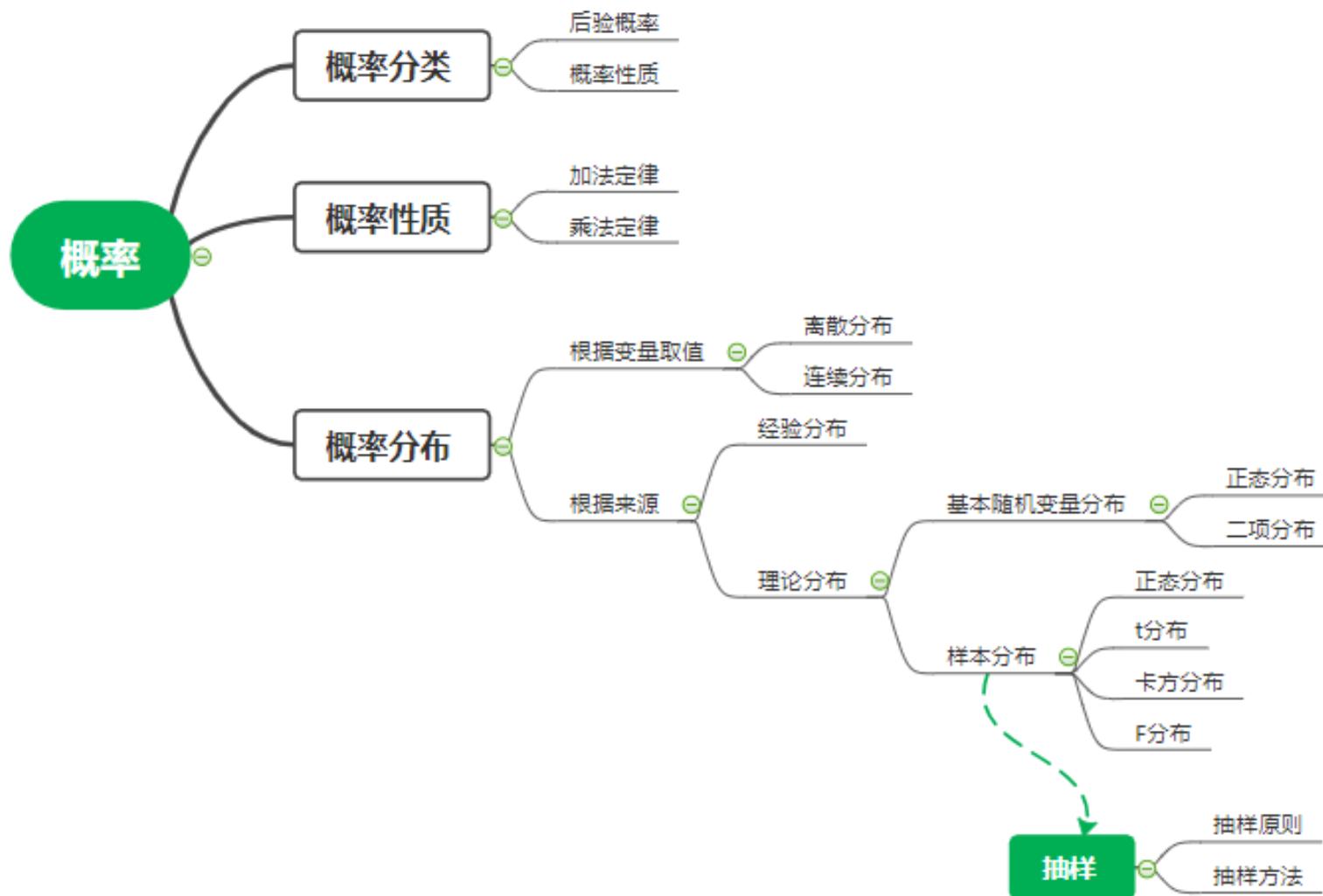


## 概览

- ◆ 概率 (单选)
- ◆ 正态分布 (选择、名词解释、简答)
- ◆ 二项分布 (单选、名词解释、简答)
- ◆ 样本分布 (选择、名词解释、简答)



# 本章结构





## 一、概率

### (一) 定义

随机事件出现可能性大小的客观指标。

简单来说，就是变量的取值在事件总体中所占的概率。

#### 1. 后验概率

在已经发生的无数次事件中，计算得到的某个事件（取值）的概率。

#### 2. 先验概率

在事件发生之前通过理论方式计算所得某个事件（取值）的概率。



## 一、概率

### (二) 基本性质

#### 1. 公理系统

- (1) 任何事件的概率 $\geq 0$
- (2) 100%发生的事件的概率=1
- (3) 100%不发生的事件的概率=0

公理(2)和(3)的逆定理不存在。

#### 2. 加法定理

两个不相容事件A和B之和（取值分组/联合）的概率等于两个事件的概率（各自取值的概率）之和。

#### 3. 乘法定律

两个独立事件同时出现的概率等于两个事件概率的乘积。



## 一、概率

### (三) 分布类型

#### 1. 离散分布和连续分布

离散分布：二项分布、泊松分布

连续分布：正态分布

#### 2. 经验分布和理论分布

经验分布：样本数据的分布

理论分布

基本随机变量分布：正态分布、二项分布

抽样分布：t分布、卡方分布、F分布



## 真题演练

高途考研 | 学习成就美好

十次硬币投掷的顺序最可能是（H表示正面朝上，T表示背面朝上）（北师大，2017）

- (1) HTHTHTHTHT
  - (2) HHTHTHTTTTH
  - (3) TTTTTTTTTTTT
- A. 顺序1最可能
  - B. 顺序2最可能
  - C. 顺序3最可能
  - D. 每个顺序概率相同



## 真题演练

高途考研 | 学习成就美好

十次硬币投掷的顺序最可能是（H表示正面朝上，T表示背面朝上）（北师大，2017）

- (1) HTHTHTHTHT
  - (2) HHTHTHTTTTH
  - (3) TTTTTTTTTTTT
- A. 顺序1最可能
  - B. 顺序2最可能
  - C. 顺序3最可能
  - D. 每个顺序概率相同

答案：D



## 真题演练

高途考研 | 学习成就美好

两个四选一的选择题，一个考生全凭猜测，两题全对的概率是（      ）。

(北师大, 2014)

- A.0.025
- B.0.0625
- C.0.50
- D.0.125



## 真题演练

高途考研 | 学习成就美好

两个四选一的选择题，一个考生全凭猜测，两题全对的概率是（ ）。

(北师大, 2014)

A.0.025                  B.0.0625

C.0.50                  D.0.125

答案：B



## 二、正态分布

### (一) 特征

1. 单峰、对称
2. 中央点最高，两侧逐渐下降，先向内弯，后向外弯，拐点在 $\pm 1$ 个标准差处，永不与横轴相交。
3. 曲线下面积为1
4. 一族分布
5. 标准差与概率面积有一定的关系

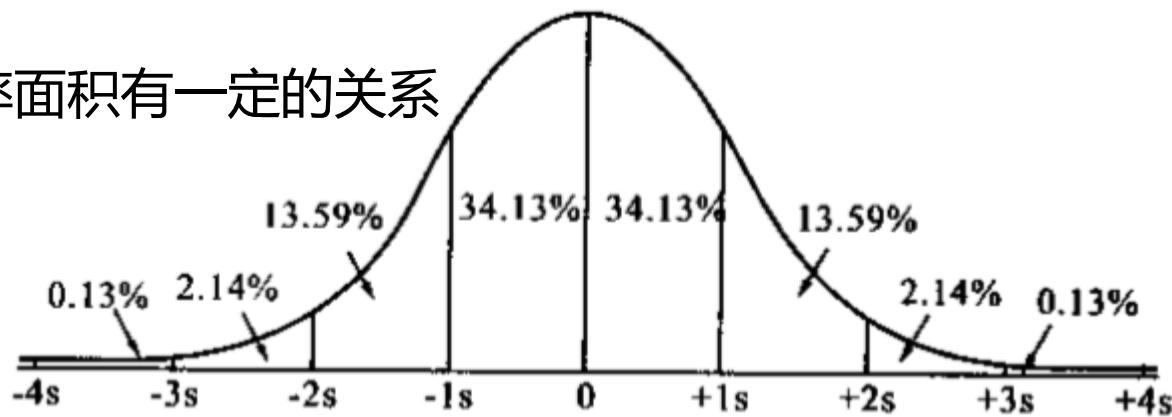


图 6-3 正态曲线下单位标准差面积比例

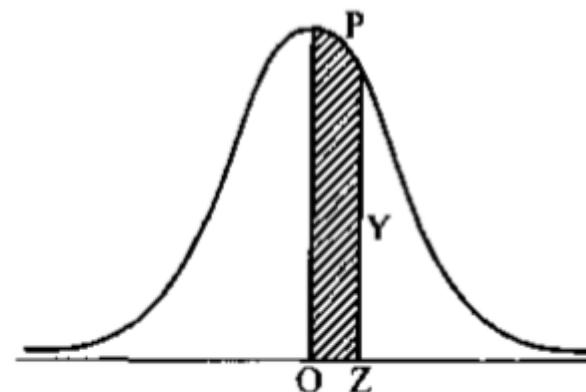


## 二、正态分布

### (二) 查表

#### 附录

附表 1 正态分布表  
(曲线下的面积与纵高)



Z	Y	P	Z	Y	P	Z	Y	P
0.00	0.39894	0.00000	0.30	0.38139	0.11791	0.60	0.33322	0.22575
0.01	0.39892	0.00399	0.31	0.38023	0.12172	0.61	0.33121	0.22907
0.02	0.39886	0.00798	0.32	0.37903	0.12552	0.62	0.32918	0.23237
0.03	0.39876	0.01197	0.33	0.37780	0.12930	0.63	0.32713	0.23565
0.04	0.39862	0.01595	0.34	0.37654	0.13307	0.64	0.32506	0.23891



## 二、正态分布

### (三) 正态检验

1. 皮尔逊偏态量数法sk

$SK=0$ 为对称分布， $SK>0$ 为正偏态分布， $SK<0$ 为负偏态分布

2. 峰度g1、偏度g2检验法

偏度系数g1： $g1=0$ 为对称分布， $g1>0$ 为正偏态分布， $g1<0$ 为负偏态分布

峰度系数g2： $g2=0$ 为正态分布的峰度， $g2>0$ 表明峰度比正态分布低阔， $g2<0$ 表明峰度比正态分布高狭

3. 累加次数曲线法



## 二、正态分布

### (四) 正态分布应用

#### 1. 把顺序数据转换成等距数据

表 6-2 3 名教师对 100 名学生的评定结果

等级	评定结果(人数)		
	教师甲	教师乙	教师丙
A	5	10	20
B	25	20	25
C	40	40	35
D	25	20	15
E	5	10	5
总数	100	100	100

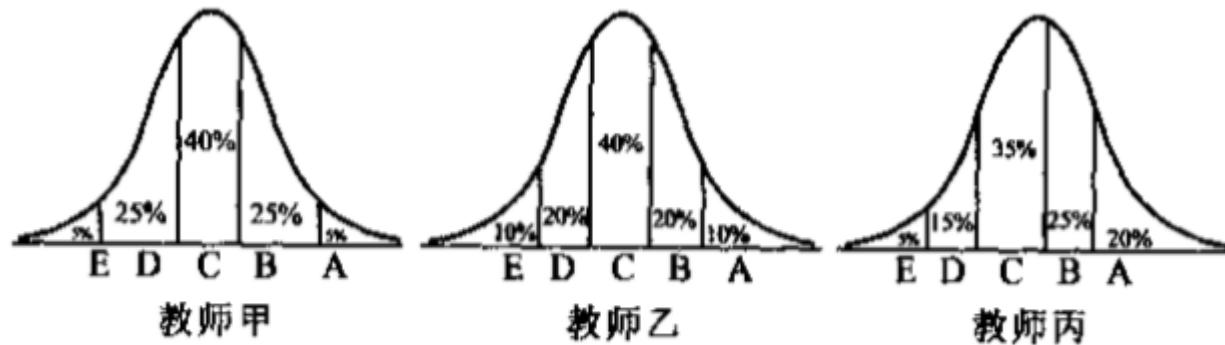


图 6-5 化等级评定为测量数据



## 二、正态分布

### (四) 正态分布应用

#### 2. 确定测验题目的难易度

表 6-5 难度分數的計算

测验题编号	通过率(%)	p 值	Z	Z + 5
1	99	0.49	-2.331	2.669
3	95	0.45	-1.645	3.355
5	85	0.35	-1.035	3.965
7	80	0.30	-0.84	4.160
9	70	0.20	-0.525	4.475
10	50	0	0	5.000
11	20	0.30	0.84	5.840
13	5	0.45	1.645	6.645
25	1	0.49	2.33	7.330



## 二、正态分布

### (四) 正态分布应用

3. 在能力分组或者等级评定时确定人数

表 6-6 能力分为 5 组时各组人數的分布

分组	各组界限	比率 p	人数分布( $p \times N$ )
A	1.8 $\sigma$ 以上	0.0359	4
B	0.6 $\sigma$ ~ 1.8 $\sigma$	0.2384	24
C	- 0.6 $\sigma$ ~ 0.6 $\sigma$	0.4514	44
D	- 1.8 $\sigma$ ~ - 0.6 $\sigma$	0.2384	24
E	- 1.8 $\sigma$ 以下	0.0359	4



## 二、正态分布

### (四) 正态分布应用

#### 4. 测验分数正态化

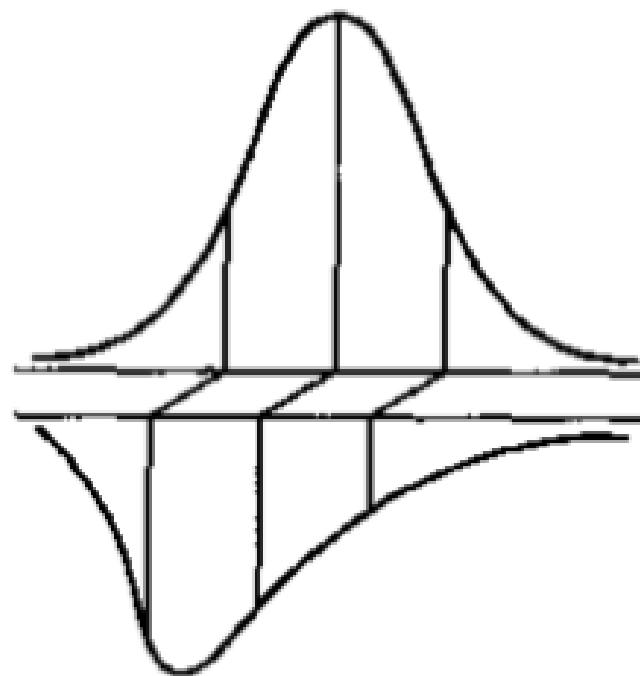


图 6-6 正态化示意图



## 真题演练

高途考研 | 学习成就美好

在标准正态曲线下，正、负三个标准差范围内的面积占总面积的比例是（ ）。 (312, 2009)

- A. 99.73%
- B. 99.00%
- C. 95.46%
- D. 95.00%

下列有关正态曲线的描述，不正确的是（ ）。 (312, 2014)

- A. 曲线以平均数为对称轴
- B.  $y$ 值越大， $P$ 值越大
- C. 曲线与x轴之间区域面积为1
- D. 曲线呈现“中间高，两边低”的形状



## 真题演练

高途考研 | 学习成就美好

在标准正态曲线下，正、负三个标准差范围内的面积占总面积的比例是

( )。 (312, 2009)

- A. 99.73%
- B. 99.00%
- C. 95.46%
- D. 95.00%

答案：A

下列有关正态曲线的描述，不正确的是 ( )。 (312, 2014)

- A. 曲线以平均数为对称轴
- B.  $y$ 值越大， $P$ 值越大
- C. 曲线与x轴之间区域面积为1
- D. 曲线呈现“中间高，两边低”的形状



## 真题演练

在标准正态曲线下，正、负三个标准差范围内的面积占总面积的比例是（ ）。 (312, 2009)

- A. 99.73%
- B. 99.00%
- C. 95.46%
- D. 95.00%

答案：A

下列有关正态曲线的描述，不正确的是（ ）。 (312, 2014)

- A. 曲线以平均数为对称轴
- B.  $y$ 值越大， $P$ 值越大
- C. 曲线与x轴之间区域面积为1
- D. 曲线呈现“中间高，两边低”的形状

答案：B



## 真题演练

高途考研 | 学习成就美好

培训机构甲让学生的数学成绩平均提高了30分，标准差为10；机构乙平均提高了40分，标准差为5。那么，更可能让学生数学成绩提高50分以上的培训机构是（      ）。 (312, 2015)

- A. 机构甲，因为它有较大的标准差
- B. 两机构相同，概率都是0.0227
- C. 机构乙，因为它提高成绩均值较大
- D. 两机构相同，概率都是0.04



## 真题演练

高途考研 | 学习成就美好

培训机构甲让学生的数学成绩平均提高了30分，标准差为10；机构乙平均提高了40分，标准差为5。那么，更可能让学生数学成绩提高50分以上的培训机构是（      ）。 (312, 2015)

- A. 机构甲，因为它有较大的标准差
- B. 两机构相同，概率都是0.0227
- C. 机构乙，因为它提高成绩均值较大
- D. 两机构相同，概率都是0.04

答案：B



## 真题演练

高途考研 | 学习成就美好

假定能力为正态分布，现取正、负各3个标准差，欲将1000名被试按照能力测验成绩等距划分为好、中、差三组。下列选项中与分配到各组的被试人数最接近的是（ ）。 (312, 2018)

- A. 199, 602, 199
- B. 179, 642, 179
- C. 159, 682, 159
- D. 139, 722, 139



## 真题演练

高途考研 | 学习成就美好

假定能力为正态分布，现取正、负各3个标准差，欲将1000名被试按照能力测验成绩等距划分为好、中、差三组。下列选项中与分配到各组的被试人数最接近的是（ ）。 (312, 2018)

- A. 199, 602, 199
- B. 179, 642, 179
- C. 159, 682, 159
- D. 139, 722, 139

答案：C



## 真题演练

高途考研 | 学习成就美好

某次高考分數呈正态分布，以此为基础可以（      ）。(多选) (312, 2009)

- A. 计算考生的标准分数
- B. 由P值，计算Z值
- C. 确定某一分数界限内的考生比例
- D. 知道计划录取人数后确定录取分数线



## 真题演练

高途考研 | 学习成就美好

某次高考分數呈正态分布，以此为基础可以（      ）。(多选) (312, 2009)

- A. 计算考生的标准分数
- B. 由P值，计算Z值
- C. 确定某一分数界限内的考生比例
- D. 知道计划录取人数后确定录取分数线

答案：ABCD



## 三、二项分布

### (一) 二项试验与二项分布

二项实验：

任何一次试验只有两个结果， $A$ 和 $A^-$ 。

共有 $n$ 次试验，并且是 $n$ 是预先给定的任一整数。

每次试验各自独立，各次试验之间无相互影响。

某种结果出现的概率在任何一次试验中都是固定的。



## 三、二项分布

### (一) 二项试验与二项分布

二项分布的定义：

设有n次试验，各次试验是彼此独立的，每次试验某事件出现的概率都是p，某事件不出现的概率都是q ( $q=1-p$ )，则对于某事件出现X次 ( $0, 1, 2, \dots, n$ ) 的概率分布为

$$b(x, n, p) = C_n^x p^n q^{n-x}$$



## 三、二项分布

### (二) 二项分布的性质

#### 1. 特点

(1) 二项分布是离散型分布，概率直方图是跃阶式。

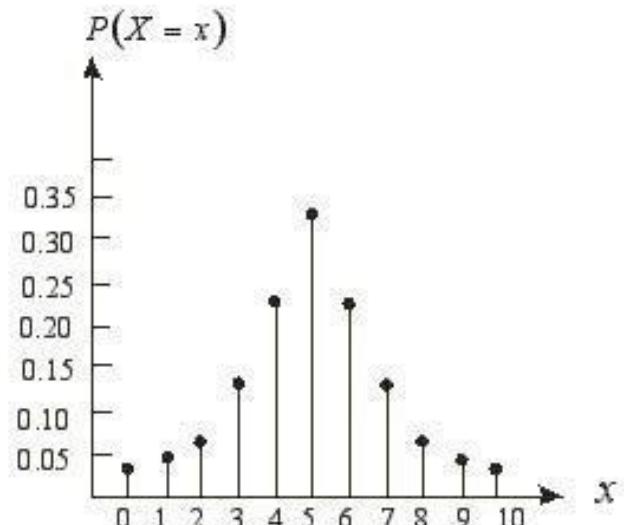
①  $p = q$  时图形对称。

②  $p \neq q$  时成偏态，若  $n$  很大，

则二项分布接近于正态分布。

(2) 若二项分布满足  $np$  且  $nq \geq 5$ ，  
二项分布接近正态分布

$$\mu = np, \sigma = \sqrt{npq}$$





## 三、二项分布

### (二) 二项分布的性质

#### 2. 应用

解决含有随机性质的问题。

例：10道判断题，答对多少道才能认为不是出于猜测？

解：已知猜对和猜错的概率是 $p=q=0.5$ ， $np=nq=5$ ，二项分布接近正态分布。

$$\mu=np=5 \quad \sigma = \sqrt{npq} = 1.58$$

当 $z=1.645$ 的时候，该点以下包含了全体的95%，也就是纯粹凭借猜测答对的概率是5%，这时候我们认为其不是出于猜测。

$$X=\mu+z^*\sigma=7.6=8$$

答对8道题，我们认为其不是纯粹猜测。



## 真题演练

高途考研 | 学习成就美好

有10道是非题，要从统计上（99%的把握）判断一个被试是否因猜测因素答对，他至少应正确回答的题目数是（      ）。（312，2013）

- A. 5
- B. 6
- C. 7
- D. 8

假设有25道五选一的选择题，随机猜测的话，猜对题数的平均值和标准差分别是（      ）。（江西师大，2016）

- A. 5, 2
- B. 2, 5
- C. 5, 5
- D. 5, 12.5



有10道是非题，要从统计上（99%的把握）判断一个被试是否因猜测因素答对，他至少应正确回答的题目数是（ ）。 (312, 2013)

- A. 5      B. 6      C. 7      D. 8

答案：D

假设有25道五选一的选择题，随机猜测的话，猜对题数的平均值和标准差分别是（ ）。(江西师大, 2016)

- A. 5, 2      B. 2, 5      C. 5, 5      D. 5, 12.5



## 真题演练

高途考研 | 学习成就美好

有10道是非题，要从统计上（99%的把握）判断一个被试是否因猜测因素答对，他至少应正确回答的题目数是（      ）。（312，2013）

- A. 5                      B. 6                      C. 7                      D. 8

答案：D

假设有25道五选一的选择题，随机猜测的话，猜对题数的平均值和标准差分别是（      ）。（江西师大，2016）

- A. 5, 2                      B. 2, 5                      C. 5, 5                      D. 5, 12.5

答案：A



## 四、样本分布

### (一) 含义

又称抽样分布，即以**样本统计量**为数据构成的分布。

- ①设存在某总体 $x_1$ 、 $x_2$ 、 $x_3$ ……
- ②从中抽取样本量为n的样本
- ③计算其样本平均数、样本标准差等统计量
- ④重复②和③无数次
- ⑤样本量为n的样本平均数和样本标准差组成的分布即为样本分布

谈论样本统计量时，首先要保证**各个样本是独立的**，各个样本都**服从同样的分布**——采用随机抽样的方法。



## 四、样本分布

### (二) 样本平均数的分布

1. 总体正态，方差( $\sigma^2$ )已知时， $\bar{X}$ 成正态分布

样本分布的平均数  $\mu_{\bar{X}} = \mu$

样本分布的方差  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

样本分布的标准差 (也称标准误SE)  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$



## 四、样本分布

### (二) 样本平均数的分布

2. 总体非正态，方差( $\sigma^2$ )已知，样本容量足够大 ( $n>30$ ) 时， $\bar{X}$ 成渐近正态分布

样本分布的平均数  $\mu_{\bar{X}} = \mu$

样本分布的方差  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

样本分布的标准差 (也称标准误SE)  $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$



## 四、样本分布

### (二) 样本平均数的分布

3. 总体正态，总体方差未知时， $\bar{X}$ 成t分布

样本分布的平均数  $\mu_{\bar{X}} = 0$

样本分布的标准差（标准误）

$$S_{\bar{X}} = \frac{s}{\sqrt{n-1}} = \frac{s_{n-1}}{\sqrt{n}} \quad s = \sqrt{\frac{\sum x^2}{n}}$$

$$s_{n-1} = \sqrt{\frac{\sum x^2}{n-1}}$$



## 四、样本分布

### (二) 样本平均数的分布

4. 总体非正态，总体方差未知，样本容量足够大 ( $n>30$ ) 时， $\bar{X}$  成近似t分布

样本分布的平均数  $\mu_{\bar{X}} = 0$

样本分布的标准差（标准误）

$$S_{\bar{X}} = \frac{s}{\sqrt{n-1}} = \frac{s_{n-1}}{\sqrt{n}} \quad s = \sqrt{\frac{\sum x^2}{n}}$$

$$s_{n-1} = \sqrt{\frac{\sum x^2}{n-1}}$$



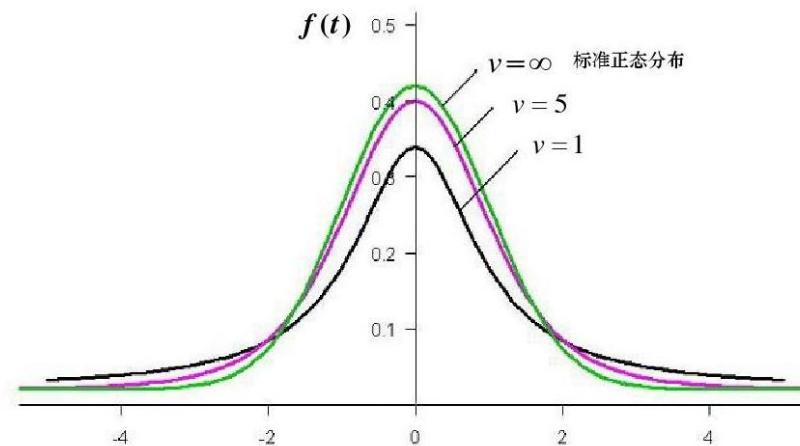
## 四、样本分布

### (三) t分布

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n-1}} \quad s = \sqrt{\frac{\sum x^2}{n}}$$

t分布形态与 $\sigma$ 无关，与n-1（自由度）有关。

t分布中，自由度df=n-1





## 四、样本分布

### (三) t分布

特点：

- ①  $\mu=0$
- ② 以平均值为0**左右对称**，左侧t为负值，右侧t为正值
- ③ 变量取值在 $-\infty \sim +\infty$ 之间
- ④ 样本容量n趋于 $\infty$ 时，t分布为正态分布，方差=1；  
当 $n-1 > 30$ 时，t分布接近正态分布，方差>1；  
当 $n-1 < 30$ 时，t分布和正态分布相差较大；  
随 $n-1$ 减少，离散程度越大，分布图中间变低尾部变高。

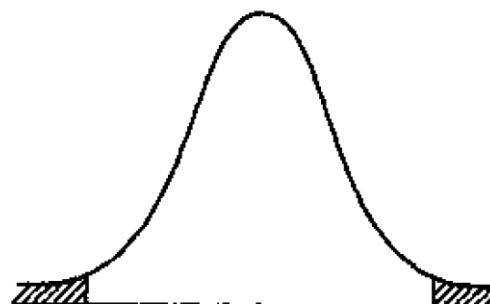


## 四、样本分布

### (三) t分布

t表

附表2 t值表



df	最大t值的概率(双侧界限)								
	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	0.765	0.978	1.250	1.638	2.351	3.182	4.541	5.841	12.941
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859



## 四、样本分布

### (四) 方差和标准差的样本分布

总体正态，样本容量足够大 ( $n > 30$ )， $s$ 、 $s^2$  成渐近正态分布。

样本分布的平均数  $\bar{X}_s = \sigma$      $\bar{X}_{s^2} = \sigma^2$

标准差分布的标准差     $\sigma_s = \frac{\sigma}{\sqrt{2n}}$

标准差分布的方差     $\sigma_s^2 = \frac{\sigma^2}{2n}$



## 四、样本分布

### (五) 卡方 ( $\chi^2$ ) 分布

① 设存在某总体  $x_1, x_2, x_3 \dots \dots$

② 从中抽取样本量为  $n$  的样本

③ 计算各个值的平方和  $\sum_{i=1}^n x_i^2$

④ 重复②和③无数次

⑤ 无数个  $\sum_{i=1}^n x_i^2$  构成的分布即为  $\chi^2$  分布

如果把原始值  $x$  换成  $z$  值，同样符合  $\chi^2$  分布

$$\chi^2 = \frac{\sum(X-\mu)^2}{\sigma^2} \quad df=n$$

如果正态总体的平均数  $\mu$  未知，用样本平均数作为估计

$$\chi^2 = \frac{\sum(X_i-\bar{X})^2}{\sigma^2} = \frac{ns^2}{\sigma^2} \quad df=n-1$$

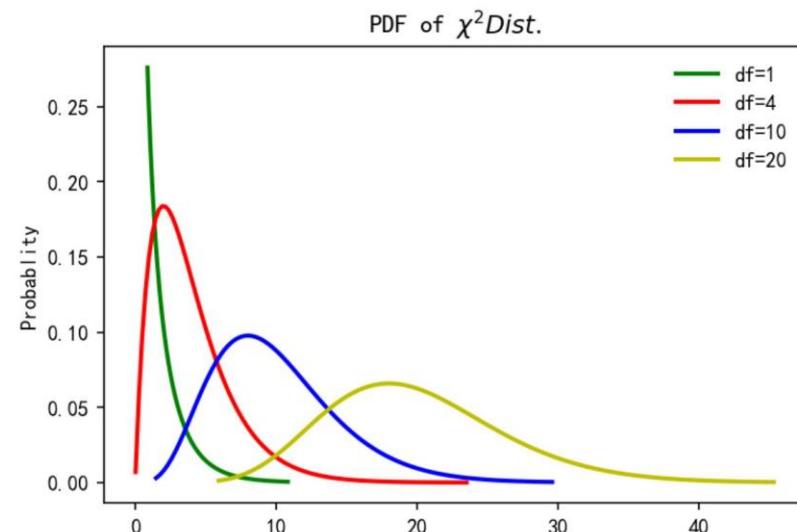


## 四、样本分布

### (五) 卡方 ( $\chi^2$ ) 分布

特点：

- 一族正偏态分布。n-1越小，分布越偏斜；df很大时，接近正态分布； $df \rightarrow \infty$ 时，即为正态分布。
- 都是正值。
- $\chi^2$ 分布的和依然是 $\chi^2$ 分布，因此 $\chi^2$ 分布具有可加性。
- 如果 $df > 2$ ， $\chi^2$ 分布的 $\mu_{\chi^2} = df$ ,  $\sigma_{\chi^2}^2 = 2df$
- $\chi^2$ 是连续型分布，但是有些离散性分布也接近 $\chi^2$ 分布。





## 四、样本分布

### (六) F分布

$\frac{\chi_1^2/\text{df}_1}{\chi_2^2/\text{df}_2}$ 被称为F比率，无数个F的分布构成F分布。

方差分析中F检验就基于此。

$$F = \frac{\chi_1^2/\text{df}_1}{\chi_2^2/\text{df}_2} = \frac{S_{n_1-1}^2/\sigma_1^2}{S_{n_2-1}^2/\sigma_2^2}$$

如果抽样来自一个总体， $\sigma_1^2 = \sigma_2^2$

$$F = \frac{S_{n_1-1}^2}{S_{n_2-1}^2} \quad \text{分子自由度 df}_1 = n_1 - 1 \quad \text{分母自由度 df}_2 = n_2 - 1$$



## 四、样本分布

### (六) F分布

特点：

- 正偏态分布；  
    分布曲线随着分子分母自由度的不同而变化；  
    随着df1和df2的增加渐趋正态分布。
- 总是正值。
- 分子自由度为1时，F值与分母自由度相同的概率的t值（双侧概率）的平方相等。

$$F_{0.05} (1, 20) = 4.35 = 2.086^2 = t_{0.05/2}^2$$



## 真题演练

高途考研 | 学习成就美好

总体服从正态分布且方差已知时，其样本平均数的分布是（ ）。

(312, 2010)

- A.  $\chi^2$ 分布
- B. t分布
- C. F分布
- D. 正态分布

下列关于t分布的表述，错误的是（ ）。 (312, 2011)

- A. 对称分布
- B. 随着n的大小而变化的一簇曲线
- C. 自由度较小时，分布是均匀分布
- D. 自由度越大，t分布越接近标准正态分布



## 真题演练

高途考研 | 学习成就美好

总体服从正态分布且方差已知时，其样本平均数的分布是（ ）。

(312, 2010)

- A.  $\chi^2$ 分布
- B. t分布
- C. F分布
- D. 正态分布

答案：D

下列关于t分布的表述，错误的是（ ）。 (312, 2011)

- A. 对称分布
- B. 随着n的大小而变化的一簇曲线
- C. 自由度较小时，分布是均匀分布
- D. 自由度越大，t分布越接近标准正态分布



## 真题演练

高途考研 | 学习成就美好

总体服从正态分布且方差已知时，其样本平均数的分布是（      ）。

(312, 2010)

- A.  $\chi^2$ 分布
- B. t分布
- C. F分布
- D. 正态分布

答案：D

下列关于t分布的表述，错误的是（      ）。 (312, 2011)

- A. 对称分布
- B. 随着n的大小而变化的一簇曲线
- C. 自由度较小时，分布是均匀分布
- D. 自由度越大，t分布越接近标准正态分布

答案：C



## 真题演练

高途考研 | 学习成就美好

下列关于 $\chi^2$ 分布的表述，正确的是（ ）。(312, 2009)

- A.  $\chi^2$ 取值永远不会小于0
- B. 其均值等于其自由度
- C. 随着自由度的增大， $\chi^2$ 分布趋于正偏态
- D. 多个标准正态分布变量的线性组合所得的新变量符合 $\chi^2$ 分布

样本平均数抽样分布趋向于正态分布的必要条件是（ ）。(312, 2015)

- A. 总体分布单峰、对称
- B. 总体均值、方差已知
- C. 总体分布不限，大样本
- D. 总体分布正态，样本方差已知



下列关于 $\chi^2$ 分布的表述，正确的是（ ）。 (312, 2009)

- A.  $\chi^2$ 取值永远不会小于0
- B. 其均值等于其自由度
- C. 随着自由度的增大， $\chi^2$ 分布趋于正偏态
- D. 多个标准正态分布变量的线性组合所得的新变量符合 $\chi^2$ 分布

答案：A

样本平均数抽样分布趋向于正态分布的必要条件是（ ）。 (312, 2015)

- A. 总体分布单峰、对称
- B. 总体均值、方差已知
- C. 总体分布不限，大样本
- D. 总体分布正态，样本方差已知



下列关于 $\chi^2$ 分布的表述，正确的是（ ）。 (312, 2009)

- A.  $\chi^2$ 取值永远不会小于0
- B. 其均值等于其自由度
- C. 随着自由度的增大， $\chi^2$ 分布趋于正偏态
- D. 多个标准正态分布变量的线性组合所得的新变量符合 $\chi^2$ 分布

答案：A

样本平均数抽样分布趋向于正态分布的必要条件是（ ）。 (312, 2015)

- A. 总体分布单峰、对称
- B. 总体均值、方差已知
- C. 总体分布不限，大样本
- D. 总体分布正态，样本方差已知

答案：C



# 第七章 抽样原理及方法



## 概览

- ◆ 抽样原则（单选）
- ◆ 抽样方法（选择）



## 本章结构





## 一、抽样原则

### (一) 抽样原则

#### 基本原则：随机性原则

- 每个个体被选取的概率相等。
- 进行返回取样，以保证每个个体每次被抽取的概率不变。

目的：有很大的可能性使得样本保持和总体相同的结构，即可以保证样本代表总体。

其他原则：代表性原则、置换性原则



## 一、抽样原则

### (二) 样本容量

1. 样本容量与总体没有固定的数量关系；
2. 一般来说，样本容量越大，产生的误差越小；
3. 若总体差异很大，则需要增大样本容量以减少抽样误差；
4. 在已知置信区间和置信水平的前提下，可以通过公式和查表的方式计算或找到对应的样本量大小。



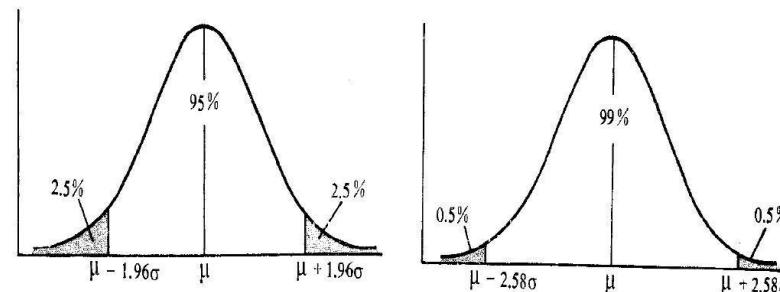
## 一、抽样原则

### (三) 样本容量计算

抽样误差:  $\bar{X} - \mu$

最大允许抽样误差  $d = Z \times SE_{\bar{x}}$

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$





## 真题演练

高途考研 | 学习成就美好

抽样的基本原则是（ ）。 (312, 2008)

- A. 随机化原则
- B. 标准化原则
- C. 概括化原则
- D. 等距化原则

用简单随机抽样方法抽取样本，如果要使抽样标准误降低50%，则样本容量需扩大的倍数是（ ）。 (312, 2012)

- A. 2
- B. 4
- C. 5
- D. 8



## 真题演练

高途考研 | 学习成就美好

抽样的基本原则是（ ）。(312, 2008)

- A. 随机化原则
- B. 标准化原则
- C. 概括化原则
- D. 等距化原则

答案：A

用简单随机抽样方法抽取样本，如果要使抽样标准误降低50%，则样本容量需扩大的倍数是（ ）。(312, 2012)

- A. 2
- B. 4
- C. 5
- D. 8



## 真题演练

高途考研 | 学习成就美好

抽样的基本原则是（ ）。(312, 2008)

- A. 随机化原则
- B. 标准化原则
- C. 概括化原则
- D. 等距化原则

答案：A

用简单随机抽样方法抽取样本，如果要使抽样标准误降低50%，则样本容量需扩大的倍数是（ ）。(312, 2012)

- A. 2
- B. 4
- C. 5
- D. 8

答案：B



## 二、抽样方法

### （一）简单随机取样法

包括抽签法，随机数字法。

评价：

适合总体数目较小，个体差异较小时用；

最能体现随机化原则；

在大规模抽样时费时费力；

容易忽略总体已有信息。



## 二、抽样方法

### (二) 等距取样法

又称系统抽样，机械抽样。对样本进行排序，之后每间隔一定距离抽取一个样本。

评价：

适合总体数目庞大时用；

抽样方法简单，样本代表性比简单随机抽样好；

总体具有周期性变化时代表性不如简单随机抽样；

可能忽略已有信息。



## 二、抽样方法

### (三) 分层随机取样法

方法：按照总体已有的某些特征，将总体分成几个不同的部分（每个部分叫一个层），再分别在每一层中随机抽样。

两种方式：

- (1) 按各层人数比例分配：人多的层多分配，人少的层少分配。
- (2) 最佳分配：标准差大的层多分配，标准差少的人少分配。

评价：

个体差异较大时用；

充分利用了总体已知的信息；

样本代表性及推论精确性优于简单随机抽样；

层内尽量同质，层间尽量异质，层间差异越大，分层效果越好。



## 二、抽样方法

### (四) 多段随机取样法

如两阶段随机取样，总体容量很大时用。

方法：

将总体分层M个集团，从中抽取m个作为第一阶段样本；

分别从m个集团中抽取个体 ( $n_i$ ) 构成第二阶段样本。

评价：

相对于简单随机抽样，标准误大一些。

抽样简便易行，节省经费。



## 真题演练

当总体中的个体分布具有周期性规律时，不适用的抽样方法是（ ）。  
(312, 2015)

- A. 简单随机抽样
- B. 分层抽样
- C. 等距抽样
- D. 方便抽样

关于分层随机抽样的特点，表述正确的是（ ）。 (312, 2011)

- A. 总体中的个体被抽取的概率相同
- B. 所有被试在一个层内抽取
- C. 层间异质，层内同质
- D. 层间变异小于层内变异



## 真题演练

高途考研 | 学习成就美好

当总体中的个体分布具有周期性规律时，不适用的抽样方法（      ）。

(312, 2015)

- A. 简单随机抽样
- B. 分层抽样
- C. 等距抽样
- D. 方便抽样

答案：C

关于分层随机抽样的特点，表述正确的是（      ）。 (312, 2011)

- A. 总体中的个体被抽取的概率相同
- B. 所有被试在一个层内抽取
- C. 层间异质，层内同质
- D. 层间变异小于层内变异



## 真题演练

当总体中的个体分布具有周期性规律时，不适用的抽样方法是（ ）。  
(312, 2015)

- A. 简单随机抽样
- B. 分层抽样
- C. 等距抽样
- D. 方便抽样

答案：C

关于分层随机抽样的特点，表述正确的是（ ）。 (312, 2011)

- A. 总体中的个体被抽取的概率相同
- B. 所有被试在一个层内抽取
- C. 层间异质，层内同质
- D. 层间变异小于层内变异

答案：C



## 真题演练

高途考研 | 学习成就美好

某企业想考察不同工龄员工的离职意向，将整个企业3000名员工按工龄分为三组：两年以下有1000人，两年至五年有1500人，五年以上有500人。如果采用按比例分层随机抽样的方法抽取容量为600的样本，那么从两年至五年工龄的员工中应抽取的人数为（      ）。 (312, 2017)

- A. 100
- B. 200
- C. 300
- D. 400



## 真题演练

某企业想考察不同工龄员工的离职意向，将整个企业3000名员工按工龄分为三组：两年以下有1000人，两年至五年有1500人，五年以上有500人。如果采用按比例分层随机抽样的方法抽取容量为600的样本，那么从两年至五年工龄的员工中应抽取的人数为（      ）。(312, 2017)

- A. 100
- B. 200
- C. 300
- D. 400

答案：C

# THANKS

- 期待下次相遇 -



高途学院APP下载



高途学院公众号

