点睛班

同等学力国考管理科学与工程综合

统计学

主讲老师：刘老师

统计学

第一章 数据的整理与图形表示
第二章 随机变量以及抽样分布
第三章 参数估计
第四章 假设检验
第五章 线性回归分析
第六章 时间序列分析

点睛 讲解内容

一、 各章重难点
二、 常见题型、解答策略

第一章 经济学导论——考点总结

序号	考点	题型
1	图形的适用	选择、简答
2	频数分布表的绘制（等距分组）	分析计算题
3	**数据的描述性指标的比较、计算**	选择、简答、分析计算

考点1：各种图形的适用

- 数据类型
  - 品质数据
    - 汇总表
      - 条形图
      - 饼图
      - 环形图
  - 数值型数据
    - 原始数据
      - 茎叶图
      - 箱线图
    - 分组数据
      - 直方图
      - 折线图
    - 时序数据
      - 线图
    - 多元数据
      - 散点图
      - 气泡图
      - 雷达图

小贴士：✷数据也有级别高低之分。定类→定序→定量；

✷高级的数据适用低级数据的方法，反之不然。

读研教育
www.duyan.com

考点2：频数分布表的绘制

绘制直方图，要先编制频数分布表

编制组距数列的步骤

排序，确定最大值、最小值，计算全距（也叫极差）

确定组距数列类型。等距OR不等距？

确定组数和组距。

确定组限和组限表示方法。同限OR异限？

计算频数、频率，编制统计表

考点3：数据的描述性指标的比较、计算

序号	考点	题型
集中趋势指标	众数的优缺点	选择、简答
	中位数的计算、优缺点	
	均值的计算，优缺点	
离中趋势指标	极差	
	标准差	

考点3：数据的描述性指标的比较、计算

成绩（分）	人数（人）f	组中值 x	每组总成绩（分）xf
60-70	10	65	650
70-80	20	75	1500
80-90	15	85	1275
90以上	5	95	475
合计	50	-	3900

$$\bar{x} = \frac{\sum xf}{\sum f}$$

$$\bar{x} = \sum x \cdot \frac{f}{\sum f}$$

考点3：数据的描述性指标的比较、计算

- 算术平均数：
  - 易受极端值影响(使用了全部数据)
  - 数学性质优良,主要用于数值型数据
  - 数据对称分布或接近对称分布时应用
- 中位数：
  - 不受极端值影响
  - 数据分布偏斜程度较大时应用;主要用于顺序数据
- 众数：
  - 不受极端值影响
  - 数据分布偏斜程度较大时应用;主要用于分类数据

考点3：数据的描述性指标的比较、计算

总体标准差与样本标准差

	总体方差	样本方差
未分组数据	$\sigma^2=\dfrac{\sum_{i=1}^{N}(X_i-\overline{X})^2}{N}$	$s^2=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}$
分组数据	$\sigma^2=\dfrac{\sum_{i=1}^{K}(X_i-\overline{X})^2 f_i}{\sum_{i=1}^{K} f_i}$	$s^2=\dfrac{\sum_{i=1}^{k}(x_i-\overline{x})^2 f_i}{\sum_{i=1}^{k} f_i-1}$

★能综合反映全部单位标志值的实际差异程度；

★用平方的方法消除各标志值与算术平均数离差的正负值问题，可方便地用于数学处理和统计分析运算.

★标准差越大，说明数据越离散，数据之前的差异越大。

第二章 随机变量及抽样分布

序号	考点	题型
1	正态分布的性质、概率计算	简答、分析计算
2	二项分布总体均值与标准差	简答
3	四种抽样方法的特点	选择、简答
4	样本均值的分布	是抽样推断的基础知识

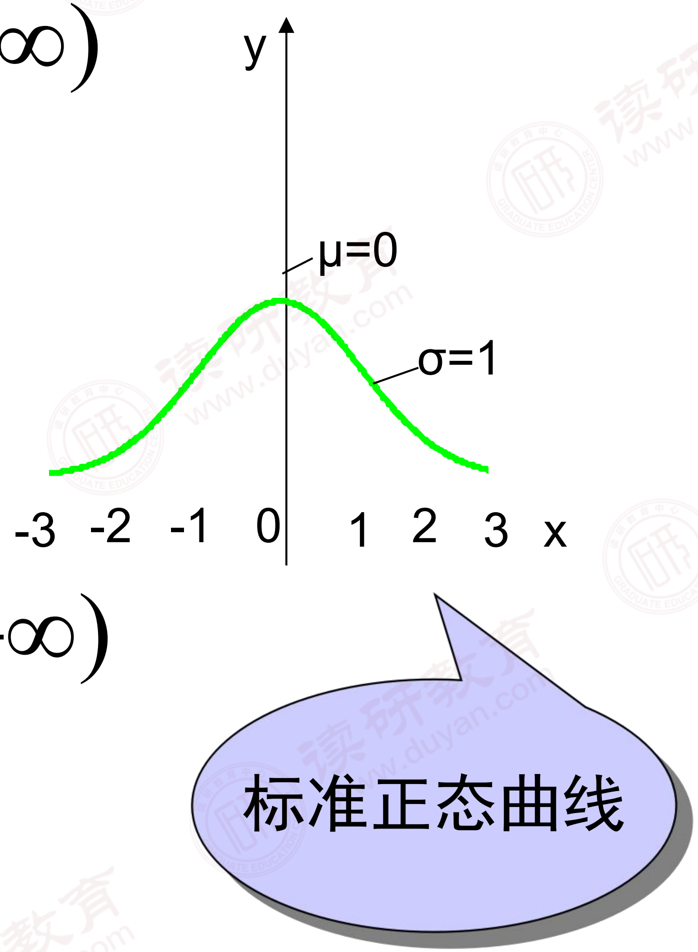
考点1：正态分布函数的性质、概率计算

正态总体的函数表示式

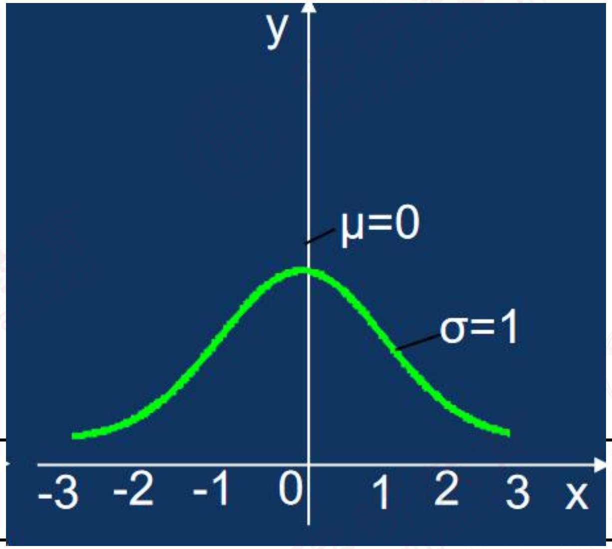
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x\in(-\infty,+\infty)$$

当μ= 0，σ=1时

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \qquad x\in(-\infty,+\infty)$$



考点1：正态分布函数的性质、概率计算

<table>
<tr><th colspan="5">正态分布函数的性质</th></tr>
<tr><th>类型</th><th>表达式</th><th>图形</th><th colspan="2">性质</th></tr>
<tr><td>正态分布</td><td>$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$</td><td rowspan="4">
</td><td>1</td><td>曲线在x轴的上方，与x轴不相交.</td></tr>
<tr><td rowspan="3">标准正态分布</td><td rowspan="3">$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$</td><td>2</td><td>曲线是单峰的,它关于直线x=μ对称.</td></tr>
<tr><td>3</td><td>曲线在x=μ处达到峰值(最高点)</td></tr>
<tr><td>4</td><td>曲线与x轴之间的面积为1</td></tr>
<tr><td rowspan="2">性质</td><td>5</td><td colspan="3">当 x<μ时,曲线上升;当x>μ时,曲线下降.并且当曲线向左、右两边无限延伸时,以x轴为渐近线,向它无限靠近.</td></tr>
<tr><td>6</td><td colspan="3">当μ一定时，曲线的形状由σ确定 .σ越大，曲线越“矮胖”，表示总体的分布越分散；σ越小，曲线越“瘦高”，表示总体的分布越集中.</td></tr>
</table>

考点1：正态分布函数的性质、概率计算

首先，将随机变量 x 标准化，令 $Z=\frac{x-\mu}{\sigma}$ ，

定理：X~N (μ,σ^2)，则**Z~ N** $(0,1)$

再查表求概率

1.（15 分）假设某饮料罐装生产线罐装后饮料重量服从正态分布，均值为 10，标准差为 4。现随机从该生产线抽取一瓶测量重量，请计算该瓶重量小于 8 的概率。

根据题意有，　$X\sim N(10,4^2)$,

则 $P(x<8)=p\left(\frac{x-10}{4}<\frac{8-10}{4}\right)$

$=p(z<-0.5)=\phi(-0.5)=1-\phi(0.5)$

因为 $\phi(0.5)=0.6915$ ，

所以 $1-\phi(0.5)=0.3085$

考点2：二项分布的总体均值与标准差

二项分布的总体平均数： $\mu_x = np$

- 二项分布的总体标准差 $\sigma_x = \sqrt{npq}$

二项成数（百分数）分布的总体平均数： $\mu_p = p$

- 二项成数（百分数）分布的总体标准差： $\sigma_p = \sqrt{\dfrac{pq}{n}}$

成数的分布：在推断统计里的合格率的推断统计会用到

考点3：四种抽样方法的特点

	抽样过程	特点
简单随机抽样	从总体N个单位(元素)中随机地抽取n个单位作为样本，使得总体中每一个元素都有相同的机会（概率）被抽中	总体中每一个元素都有相同的机会（概率）被抽中
分层抽样	先分层，再从每一层中随机抽	每层中的每个元素被抽的概率相同
整群抽样	先分群（组），再随机抽群	被抽到的群中所有元素被调查
系统抽样	先从数字1到k之间随机抽取一个数字r作为初始单位，以后依次取$r+k$，$r+2k$…等单位	距离内的元素被抽的概率相同
常见题型	选择题。根据给定情形，判断适用那种抽样方法 简答题。	

考点4：样本均值的分布

- 总体分布
  - 正态分布
    - 样本均值 正态分布
  - 非正态分布
    - 样本均值 正态分布
    - 样本均值 非正态分布

读研教育 www.duyan.com

考点4：样本均值的分布

这是做抽样推断分析计算题必用的知识点

样本均值	$\bar{x}=\dfrac{\sum_{i=1}^{n}x_i}{n}$
样本方差	$s^2=\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$
样本标准差	$s=\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

$$\bar{x}\sim N\left(\mu,\frac{\sigma^2}{n}\right) \rightarrow \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

$$E(\bar{x})=\mu \qquad \sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$$

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}} \qquad \hat{\sigma}_{\bar{x}}=\frac{s}{\sqrt{n}}$$

考点4：样本均值的分布

服从自由度为 df = n - 1 的 t 分布。即：

$$t=\frac{\overline{x}-\mu}{s_{\overline{x}}}=\frac{\overline{x}-\mu}{s/\sqrt{n}}$$

$s_{\overline{x}}$ 为样本平均数的标准误，其计算公式为：

$$s_{\overline{x}}=\frac{s}{\sqrt{n}}$$

第三章 参数估计

序号	考点	题型
1	区间估计的求法	分析计算
2	样本容量的确定	选择、简答

考点1：区间估计的求法

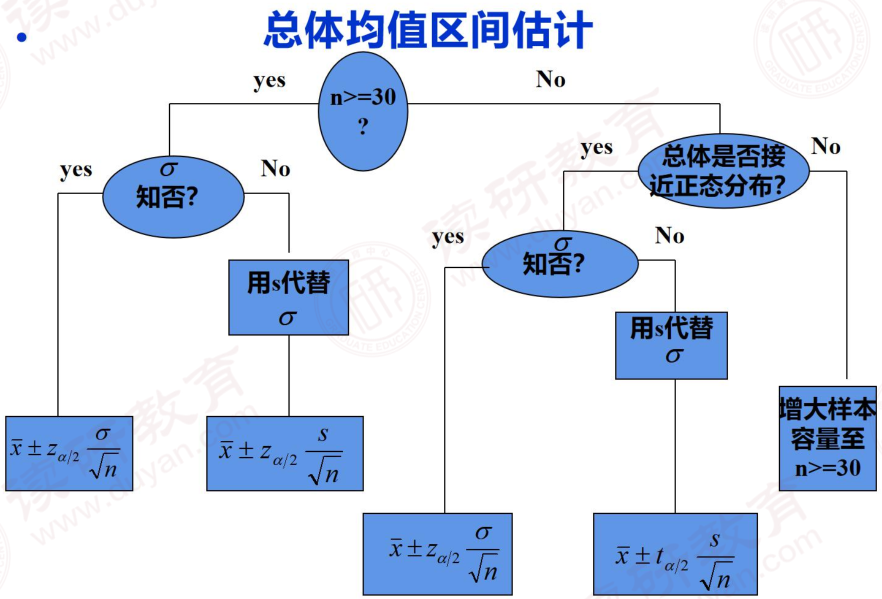
样本均值±分位数值×样本均值的标准误差

第一步，观察总体是否服从正态分布

第二步，计算样本均值、样本标准差

第三步，找置信水平1-a

第四步，计算抽样极限误差



2021-3-20

考点1：区间估计的求法

区间估计需要用到抽样极限误差

抽样平均误差（重复抽样）

抽样极限误差

$$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} \Rightarrow \Delta_{\bar{x}} = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}} \Rightarrow \Delta_{\bar{x}} = Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

考点1：区间估计的求法

总体均值在置信水平下的置信区间可一般性地表达为 ：

样本均值±分位数值×样本均值的标准误差

Z分布： $Z=\dfrac{\overline{X}-\mu}{\sigma_{\overline{X}}}$ 置信区间： $\left(\bar{x}-Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{x}+Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}\right)$

t分布： $t=\dfrac{\overline{X}-\mu}{\sigma_{\overline{X}}}$ 置信区间： $\left(\bar{x}-t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}},\ \bar{x}+t_{\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}\right)$

样本比例±分位数值×样本比例的标准误差

$$p\pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

考点1：区间估计的求法

总体参数	符号表示	样本统计量
均值差	$\mu_1-\mu_2$	$\bar{x}_1-\bar{x}_2$
比例差	$\pi_1-\pi_2$	p_1-p_2
方差比	σ_1^2/σ_2^2	s_1^2/s_2^2

两个总体均值之差(μ_1-μ_2)在置信水平下的置信区间可一般性地表达为

$(\bar{x}_1-\bar{x}_2)$±分位数值×标准误差

考点2：样本容量的确定

估计总体均值必须的样本量

在 $1-\alpha$ 的置信度下估计总体均值的允许误差为 $\Delta_{\bar{x}}$ 则必要的样本量为：

•重复抽样下：

$$\boxed{n = \frac{Z_{\alpha/2}^{2}\sigma^{2}}{(\Delta_{\bar{x}})^{2}}}$$

选择题、简答题。样本容量的确定方法：

（1）写出抽样极限误差的公式

(2)变形，观察变形后的公式就可以选出答案或计算出确切的容量。

注：关于增加还减少样本容量的选择题，之间根据右边的公式就可观察出二则的关系

$$\Delta_{\bar{x}} = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

第四章 假设检验

序号	考点	题型
1	假设检验中的两类错误	简答
2	假设检验步骤、分析计算	简答、分析计算
3	假设检验与区间估计的关系	简答

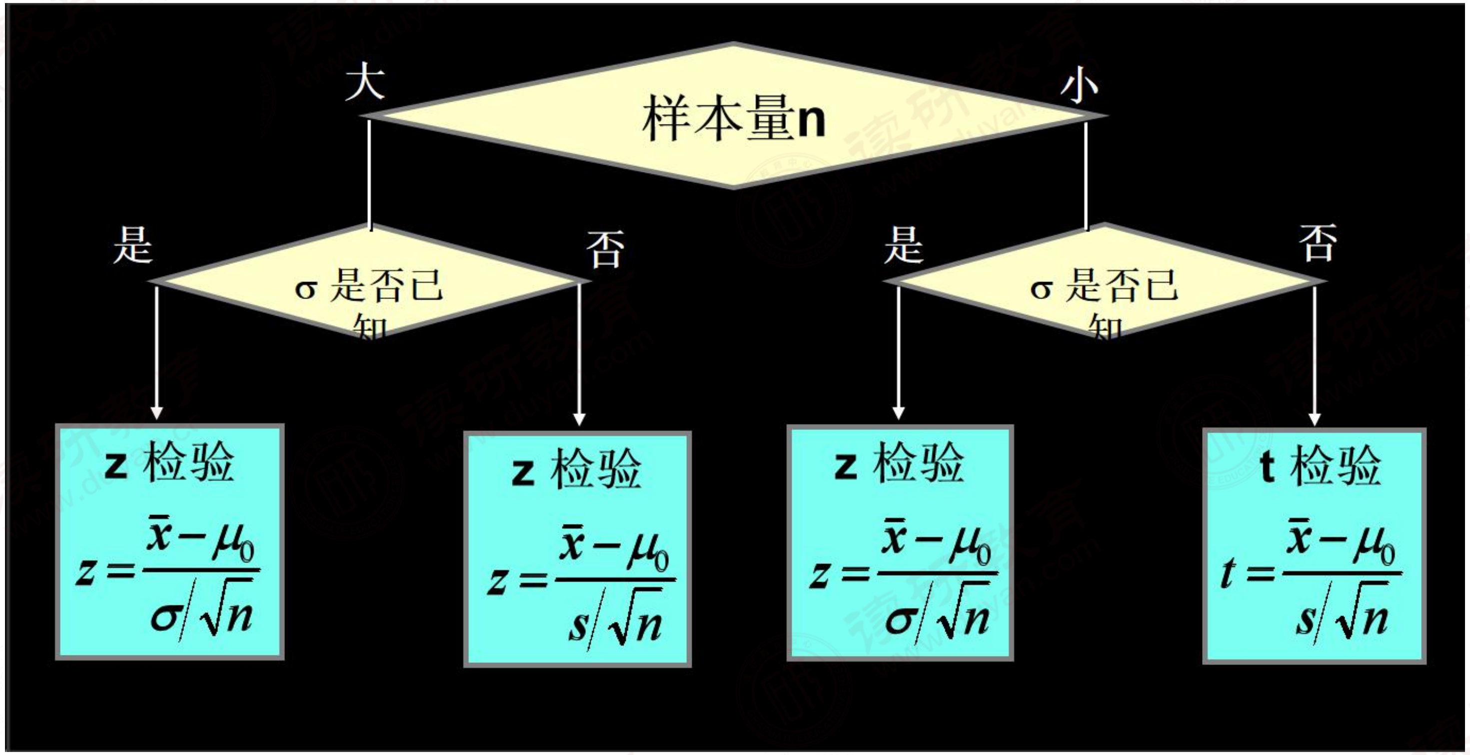
考点1：假设检验中的两类错误

- 第Ⅰ类错误(α错误)：弃真错误
  - 原假设为正确时拒绝原假设
  - 第Ⅰ类错误的概率记为 ，被称为显著性水平
- 第Ⅱ类错误(β错误)：取伪错误
  - 原假设为错误时未拒绝原假设
  - 第Ⅱ类错误的概率记为 (Beta)

考点2：假设检验步骤、分析计算

熟练假设检验的步骤，并会做计算题			
1	提出假设	原假设H_0: 想反对的	，≥，≤
		备择假设H_1: 想支持的	≠，<,>
2	给定显著性水平a以及样本容量n，确定临界值	单侧检验	Z_{α}， t_{α}
		双侧检验	$Z_{\alpha/2}, t_{\alpha/2}$
3	计算统计量	Z统计量	$z=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$
		t统计量（正态总体小样本，方差未知）	$t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\sim t(n-1)$
4	比较，并做决策	|统计量| > 临界值	拒绝H_0

考点2：假设检验步骤、分析计算



考点2：假设检验步骤、分析计算

举例：小样本，方差 已知

【例】某市历年来对7岁男孩的统计资料表明，他们的身高服从均值为1.32米、标准差为0.12米的正态分布。现从各个学校随机抽取25个7岁男学生，测得他们平均身高1.36米，若已知今年全市7岁男孩身高的标准差仍为0.12米，问与历年7岁男孩的身高相比是否有显著差异(取 α ＝0.05)。

解：从题意可知，$\bar{X}$＝1.36米，μ＝1.32米，σ＝0.12米。

(1)建立假设：H_0：$\mu=1.32$，H_1：$\mu\neq1.32$

(2)确定统计量：

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}=\frac{1.36-1.32}{0.12/\sqrt{25}}=1.67$$

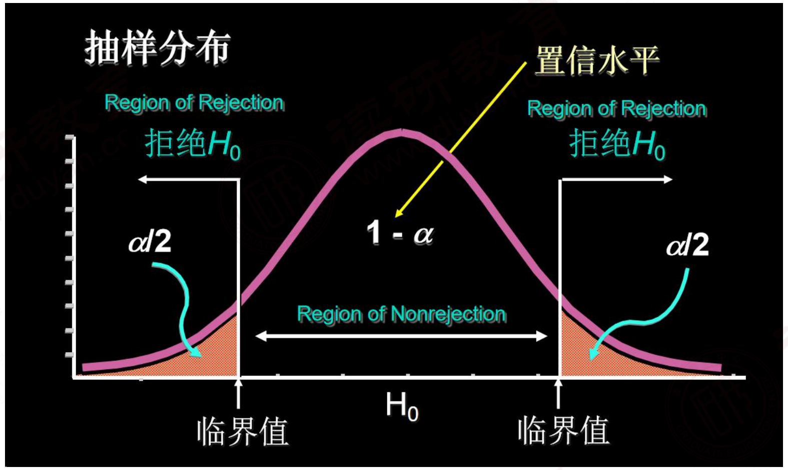
(3)比较统计量 α ＝0.05, $Z_{\alpha/2}$=1.96, $Z=1.67<1.96$,不拒绝原假设。表明没有证据表明今年7岁男孩平均身高与历年7岁男孩平均身高有显著差异

考点3：假设检验与区间估计的关系

区间估计

总体均值 μ 在1-α 置信水平下的置信区间为

$$\bar{x}\pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\ \text{或}\ \bar{x}\pm z_{\alpha/2}\frac{s}{\sqrt{n}}(\sigma\ \text{未知})$$



假设检验

|统计量| > 临界值，拒绝H_0

接受域

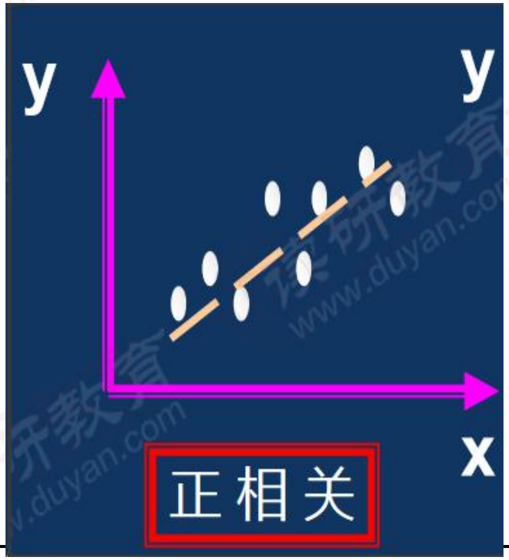
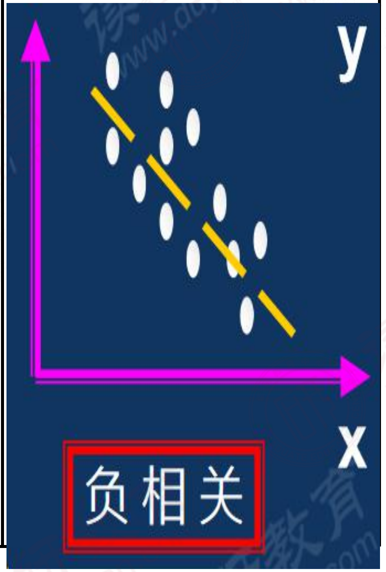
$$\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\langle\mu_0\langle\bar{x}+z_{\alpha/2}\frac{s}{\sqrt{n}})$$

可以写出区间估计的一般表达式，结合表达式来说明二者的关系

第五章 线性回归分析

序号	考点	题型
1	相关关系与回归	选择
2	解读回归估计结果	分析计算题
3	方差分析	分析计算题

考点1：相关关系与回归关系

相关关系 $-1 \leqslant r \leqslant 1$，相关关系 $\neq$ 因果关系			
类型	r值	图形	
完全相关	$\lvert r\rvert=1$		
不相关	r=0		
存在一定相关性	＜0.3 弱相关 0.3～0.5 低度相关 0.5～0.8 显著相关 0.8～1 高度相关		

考点2：解读回归估计结果

x市城镇居民人均可支配收入和人均消费性支出之间依存关系的线性回归方程：

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i = 525.8662 + 0.7083 X_i$$

$\hat{\beta}_2 = 0.7083$ 表明，当居民人均可支配收入增加1元时，人均消费性支出将平均增长0.7083元

读研教育 www.duyan.com

考点3：方差分析

有交互效应的方差分析表，会填表

变异来源	离差平方和SS	自由度df	均方MS	F值
A因素	SSA	a-1	MSA=SSA/(a-1)	F_A=MSA/MSE
B因素	SSB	b-1	MSB=SSB/(b-1)	F_B=MSB/MSE
AB交互作用	SSAB	(a-1)(b-1)	MSAB=SSAB/(a-1)(b-1)	F_{AB}=MSAB/MSE
误　差	SSE	ab(m-1)	MSE=SSE/rs(m-1)	
合　计	SST	n-1		

考点3：方差分析

无交互效应的方差分析表

变差来源	离差平方和 SS	自由度 df	均方 MS	F值
A因素	SSA	a-1	MSA=SSA/(b-1)	F_A=MSA/MSE
B因素	SSB	b-1	MSB=SSB/(b-1)	F_B=MSB/MSE
误 差	SSE	n-a-b+1	MSE=SSE/(n-a-b+1)	
合 计	SST	n-1		

考点3：方差分析

2010年分析计算题

1.（15 分）一家鞋店零售商进行了一项研究，以确定该鞋店每天的销售量是否随方圆 1 公里范围内竞争者的数量及商店的地理位置不同而异。该公司选择了三种类型的商店：单独位于郊区的商店、位于购物中心的商店、城内的商店。这些商店方圆 1 公里内的竞争者数量不同，可以被分成四类：没有竞争者、有1个 竞争者、有2个 竞争者、有3 个及以上竞争者。对其调查数据进行方差分析，得到结果如表 1 所示。请将表格补充完整（10分 ），并在 0.05的 显著性水平下说明你的分析结论（5分 ）。

变异来源	离差平方和SS	自由度df	均方MS	F值
A因素	SSA	a-1	MSA=SSA/(a-1)	F_A=MSA/MSE
B因素	SSB	b-1	MSB=SSB/(b-1)	F_B=MSB/MSE
AB交互作用	SSAB	(a-1)(b-1)	MSAB=SSAB/(a-1)(b-1)	F_{AB}=MSAB/MSE
误 差	SSE	ab(*m*-1)	MSE=SSE/rs(*m*-1)	
合 计	SST	n-1		

表1 方 差分析表

因变量：鞋日销量

误差来源	平方和B	自由度	均方和MS	F	显著性
竞争者数	1128.032		376.011	14.910	0.000
商店位置	1619.452			32.108	0.000
竞争者数与商店位置		6	75.953	3.012	0.024
误差		24	25.219		
总计	3808.458	35			

平方和	自由度	均方和 MS	F	显著性
1128.032	3	376.011	14.910	0.000
1619.452	2	809.726	32.108	0.000
455.718	6	75.953	3.012	0.024
605.256	24	25.219	1	
3808.458	35			

第六章 时间序列

序号	考点	题型
1	趋势预测	分析计算
2	季节因素测定	分析计算

考点1：趋势预测

滑动平均法（又叫滑动平均法） $F_{t+1}=\overline{Y}_t=\dfrac{Y_{t-k+1}+Y_{t-k+2}+\cdots+Y_{t-1}+Y_t}{k}$

指数平滑预测 $F_{t+1}=\alpha Y_t+(1-\alpha)F_t$

最小二乘法——直线趋势的测定 $\hat{Y}_t=b_0+b_1t$

考点2：季节因素的测定

$$\text{季节指数 } S = \frac{\text{同月(季)平均数}}{\text{总月(季)平均数}} \times 100\%$$

季节调整

在原时序数列中消除季节因素之影响，计算公式为

$$\frac{T \times S \times C \times I}{S} = T \times C \times I$$