



八、概率抽样与非概率抽样★★★

1. 概率抽样	
简单随机抽样	有放回、不放回
分层抽样	分层抽样是指先按照某种规则把总体分为不同的层，然后在不同的层内独立、随机地抽取样本。 (要求：层内差异小，层间差异大)
系统抽样	先将总体中的所有单元按一定顺序排列，在规定范围内随机抽取一个初始单元，然后按事先规定的规则抽取其他样本单元。最简单的系统抽样：等距抽样
整群抽样	将总体中所有的基本单位按照一定规则划分为互不重叠的群，抽样时直接抽取群，对抽中的群调查其全部基本单位。(要求：群间差异小，群内差异大)
多阶段抽样	要经过两个及两个以上阶段才能抽到最终样本单位



八、概率抽样与非概率抽样★★★

2.非概率抽样【口诀】自判配方

判断抽样	在抽取样本时，调查人员依据调查目的和对调查对象情况的了解，人为确定样本单元。
方便抽样	在抽取样本时，依据方便原则，以达到最大限度降低调查成本的目的。比如“拦截式”调查，在街边或居民小区拦住行人进行调查。
自愿样本	是由自愿接受调查的单元所组成的样本。比较典型的是网上调查。
配额抽样	将总体中的各单元按一定标准划分为若干类型，将样本数额分配到各类型中，从各类型中抽取样本的方法则没有严格限制，一般采用方便抽样的方法抽取样本单元。



九、误差★★

项目	内容	
抽样误差	由于抽样的随机性造成的误差。用样本统计量估计总体参数时出现的误差。	
非抽样误差	抽样框误差	由于抽样框不完善造成的误差。
	无回答误差	调查人员没能从被调查者那里得到所需要的数据，具体有两类： ①由于随机因素造成的，如调查时被调查者不在家； ②非随机因素造成的，如被调查者不愿告诉实情而拒绝回答。
	计量误差	计量误差是指由于调查所获得的数据与其真值之间不一致造成的误差。 如，调查员在调查中有意无意地诱导被调查者，调查中的提问错误或记录答案错误，调查人员有意作弊，由于问卷的原因受访者对调查问题的理解上有偏误，受访者记忆不清，受访者提供虚假数字等。



十、估计量的性质★★★

性质	内涵
无偏性	对于不放回简单随机抽样，所有可能的样本均值取值的平均值总是等于总体均值，这就是样本均值估计量的无偏性。
有效性	在同一抽样方案下，对某一总体参数 θ ，如果有两个无偏估计量 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ ，由于样本的随机性， $\hat{\theta}_1$ 的可能样本取值较更密集在总体参数真值 θ 附近，人们会认为 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效， $\hat{\theta}_1$ 的方差小于 $\hat{\theta}_2$ 的方差。
一致性	随着样本量的增大，估计量的值如果稳定于（或收敛于）总体参数的真值，这个估计量就有一致性，可称为一致估计量。



十一、一元线性回归模型★★★

回归模型	回归方程
$Y = \beta_0 + \beta_1 X + \varepsilon$	$E(Y) = \beta_0 + \beta_1 X$
β_0和β_1为模型的参数。 ① $\beta_0 + \beta_1 X$反映了由于自变量X的变化而引起的因变量Y的线性变化。 ② 误差项 ε 是个随机变量，表示除X和Y的线性关系之外的随机因素对Y的影响，是不能由X和Y的线性关系所解释的Y的变异性。	β_0是回归直线的截距， β_1是回归直线的斜率， 表示X每变动一个单位时， $E(Y)$的变动量。



十二、回归模型的拟合效果分析★★★

1. 决定系数 R^2 （拟合优度或判定系数）。

$$R^2 = \text{可解释的变异} / \text{总变异} = \text{回归平方和} / \text{总平方和}$$

$$0 \leq R^2 \leq 1$$

越接近于1，回归模型的拟合效果越好；越接近于0，拟合效果越差。

2. 回归系数的显著性检验——t检验

验证自变量X对因变量Y是否有显著影响。

原理：反证法

（1）设原假设（自变量X对因变量Y没有影响）正确；

（2）基于的抽样分布计算一次抽样情况下得到该样本或更极端样本的概率（P值）：

如果 $P < 0.05$ ，则可以在0.05的显著性水平下拒绝原假设，认为自变量X对因变量Y有显著影响，即 $\hat{\beta}_1 \neq 0$ 。



十三、时间序列及其分类★★★

表现形式不同	内容
绝对数时间序列	① 时期序列：反映现象在一段时期内发展的结果，即“过程总量”。 ② 时点序列：反映现象在一定时点上的瞬间水平，如年底总人口数表明在各年年末时点上的人口数。
相对数时间序列	相对指标值 按时间先后顺序排列后形成的序列，如 城镇人口比重 。
平均数时间序列	平均数指标值 按时间先后顺序排列后形成的序列，如 人均国内生产总值 。



十四、时间序列的水平分析★★

时间序列的水平分析指标	时间序列的速度分析指标
发展水平 平均发展水平 增长量 平均增长量	发展速度 平均发展速度 增长速度 平均增长速度



十四、时间序列的水平分析★★

1. 平均发展水平（绝对数）			计算方法
时期序列			简单算术平均数
时点 序列	连续 时点	逐日登记、逐日排列	
		指标值发生变化才登记	加权算术平均数
	间断 时点	间断时间相等	两次简单算术平均数
		间断时间不相等	一次简单算术平均数 + 一次加权算术平均数



十四、时间序列的水平分析★★

2. 增长量与平均增长量		
增长量	逐期增长量 = 报告期水平 - 前期水平	同一时间序列中，累计增长量等于相应时期内逐期增长量之和
	累计增长量 = 报告期水平 - 固定期水平	
平均增长量	$\begin{aligned}\text{平均增长量} &= \frac{\text{逐期增长量之和}}{\text{逐期增长量个数}} \\ &= \frac{\text{累计增长量}}{\text{最末时间} - \text{最初时间}}\end{aligned}$	



十五、时间序列的速度分析★★

速度	发展速度	定基发展速度	报告期发展速度/ 最初水平	① 定基发展速度等于相应时期内各环比发展速度的连乘积。 ② 两个相邻时期定基发展速度的比率等于相应时期的环比发展速度。 【口诀】 定环乘，环定比
		环比发展速度	报告期水平/前期水平	
	增长速度	定基增长速度	报告期累计增长量/基期水平=定基发展速度-1	
		环比增长速度	报告期逐期增长量/基期水平=环比发展速度-1	
平均速度	平均发展速度		$\bar{b} = \sqrt[n]{\prod_{i=1}^n b_i} = \sqrt[n]{\frac{y_n}{y_0}}$	
	平均增长速度		平均增长速度=平均发展速度-1	



十六、时间序列的分解和预测程序★★

长期趋势 (T)	时间序列在较长一段时间内呈现的持续上升或持续下降的变动。这种趋势可能是线性的，也可能是非线性的。
季节变动 (S)	时间序列在一年内重复出现的周期性波动。例如，旅游、商品销售、农业生产、交通运输等行业的生产经营中都有明显的季节变动特征。
循环波动 (C)	时间序列呈现出的非固定长度的周期性变动。例如，经济周期有涨落相间的交替波动，周期长短不一，无固定规律
不规则波动 (I)	时间序列中除去长期趋势、季节变动和循环波动之后的随机波动。不规则波动往往是由一些偶然因素引起的，难以预测和控制，因此在时间序列预测中通常不予单独考虑。

加法模型： $Y_t = T_t + S_t + C_t + I_t$

乘法模型： $Y_t = T_t \times S_t \times C_t \times I_t$

谢谢 观看

THANK YOU