生物医学统计概论(BI148-2)

Fundamentals of Biomedical Statistics

授课：林关宁

2022 春季


哪种方差分析 — 单因素？双因素？重复测量？交互作用？

- **应该进行哪种方差分析？ 完全取决于你需要比较多少因素，以及它们的性质。简而言之：**
  1. 如果你的实验设计基于**一个预测变量**（例如：药物治疗、物种、位置等），**且至少包括两个水平**（例如：治疗A、治疗B、治疗C和对照），则进行**单因素方差分析**
  2. 如果你的实验设计基于**两个、三个或更多的预测变量**（例如：药物治疗和种群、公民身份和性别、物种和性别以及药物治疗等），**并且每个变量至少有两个水平**，则进行**双因素方差分析**
  3. 如果只有**一个预测变量**，且数据存在依赖性，这意味着**受试者在不同时间点或不同条件下进行了重复测量**（例如：所有个体在出生时、儿童期和成年期进行了测试，或所有个体在连续三次药物治疗后接受了测试），则运行**单因素重复测量方差分析**
  4. 如果你**至少有两个预测变量**，并且你的数据中存在依赖性，这意味着**受试者已经被重复测量**，在不同的时间点或不同的条件下测量（例如：所有按国籍**分组**的个人在出生时、儿童期和成年期接受测试，或所有按性别**分组**的个人在连续三次药物治疗后接受测试），则运行**双因素重复测量方差分析**

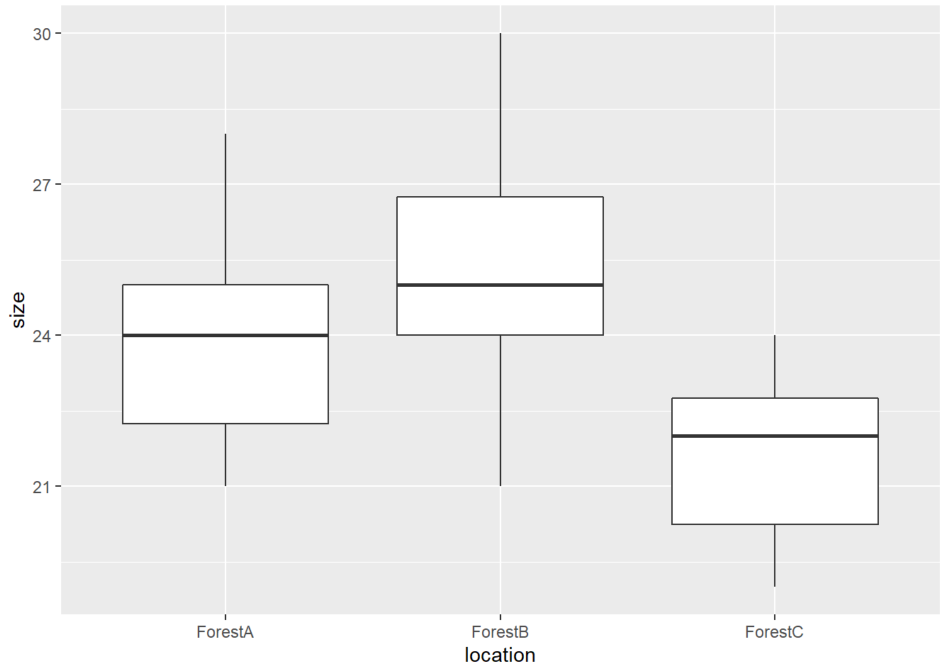
有无交互作用 →（第2条）

有无交互作用 →（第4条）

单因素方差 one-way anova

单向方差分析是一种旨在比较两个或更多样本均值的测试。零假设H0表示所有待测样本的均值相等。测试返回一个F统计量和一个p值，这将帮助您决定是否拒绝无效假设。举个例子：

假设我们想检查蓝地甲虫（Carabus indexatus）的平均大小是否因位置不同而不同。我们考虑3个不同的森林命名为A，B和C。在每个森林，我们测量10只甲虫的大小（毫米）



使用one-way anova 的假设前提是：

- 观察的独立性（每个个体仅由一个条目/测量值表示）
- 分布的正态性（对每组进行测试，例如Shapiro-Wilk测试）
- 方差的齐性/同质性（例如，用Levene检验进行检验）
- 组不包含异常值

如果我们想检查哪一组与其他组不同，需要进行一个事后测试：

- 两两t检验与功能两两 t- test，
- Tukey's Honest Significant Difference (HSD) with TukeyHSD test，
- 线性模型中的多重比较 （Bonferroni）

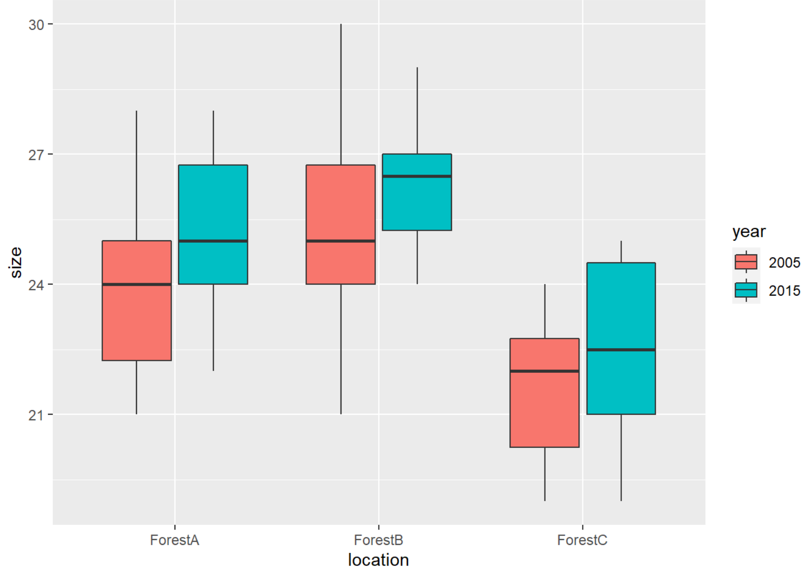
```
Analysis of Variance Table

Response: size
          Df  Sum Sq Mean Sq F value   Pr(>F)
location   2  66.467  33.233  7.1101 0.003307 **
Residuals 27 126.200   4.674
```

双因素方差 two-way anova

多因素方差分析允许基于多个自变量及其不同水平对组进行比较。因此，可以通过双向方差分析将按两个预测变量（例如位置和物种）分类的观察结果（例如树木大小、卵大小等）进行比较；如果涉及三个预测变量（例如位置、物种和年份），则可以使用三因素方差分析。

使用双因素方差分析，我们可以评估因素的主要影响 main effects，即评估每个单因素（预测变量）response variable（植株大小、卵子数量、个体体重等）的影响外，我们还可以评估看到因素之间是否存在相互作用，即特定因素的效果是否受到其他因素（其中一个）的影响。

例：我们检查蓝地甲虫的平均大小是否因位置不同而不同。我们现在引入一个新的因素：2005年和2015年在相同的3个森林A、B和C进行了测量。在每个森林，我们测量了10只甲虫的大小（以毫米为单位）（平衡设计）。总共测量了60个只。这**两个因素就是地点和年份**。



使用条件与单因素方差分析的假设基本相同：

- 观察的独立性（每个个体仅由一个条目/测量值表示）
- 分布的正态性（对每组进行测试，例如Shapiro-Wilk测试）
- 方差的同质性（例如，用Levene检验进行检验）
- 组不包含异常值

Analysis of Variance Table

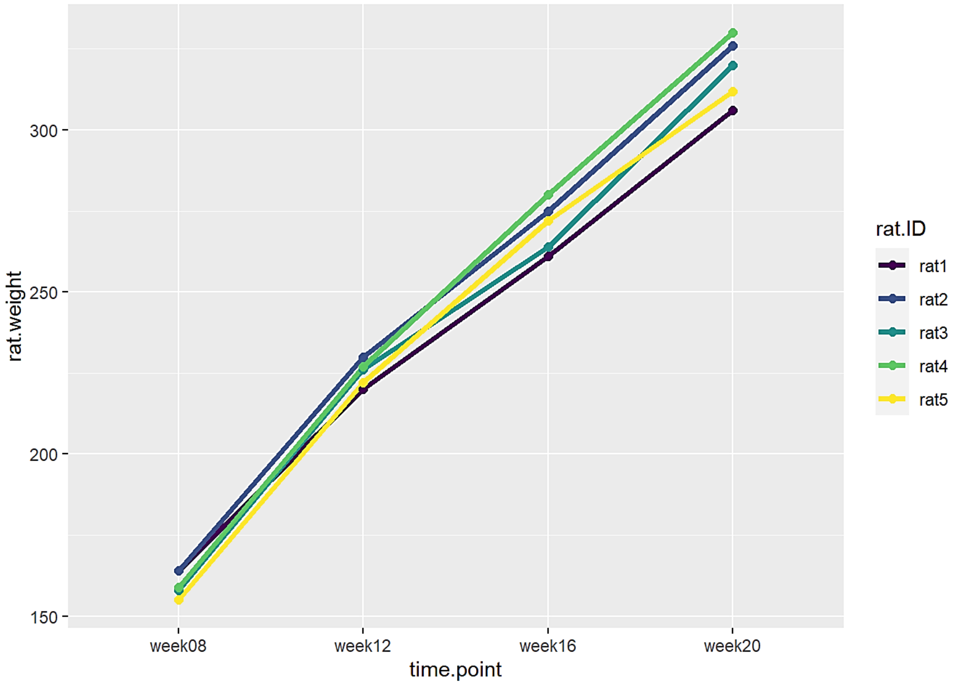
Response: size

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
location	2	146.700	73.350	17.8178	1.142e-06	***
year	1	14.017	14.017	3.4049	0.07049	.
location:year	2	0.633	0.317	0.0769	0.92606	
Residuals	54	222.300	4.117			

单因素重复测量方差 one-way repeated measure

重复测量方差分析（Repeated measures ANOVA）是一种似乎接近单因素方差分析的测试，它允许检查两组或更多组的平均值之间的差异。本质的区别在于，这些群体是相互依赖的（即相关的）。这意味着这些组包含来自相同受试者的数据或测量值。
不同是在于实验设计及数据收集的设计方面，比如在测量时间上重复：实验前、实验中、实验后或给定时间段内的每天等，或在何种情况下或条件下进行重复测量（例如，测量进行了3次，每次在应用新药后）。在该检验试中，零假设H0表示所有组的平均值相等。

举一个例子，5只大鼠称重4次，间隔4周（第8至20周）



假设是：

- 每个个体在每个测试组中以测量值的形式表示（不存在任何缺失值）
- 分布的正态性（使用Shapiro-Wilk），
- 球形检验 **sphericity** 验证，意味着在实验设计中比较任何两组（所有可能的组对）时（通常通过Mauchly 检验进行检查），验证方差齐性
- 组不包含异常值

	numDF	denDF	F-value	p-value
(Intercept)	1	12	11442.465	<.0001
time.point	3	12	793.995	<.0001

统计学之球形检验（Mauchly's test of sphericity）

球形检验（Mauchly’s test of sphericity），适用于**重复测量**时检验不同测量之间的差值的方差是否相等

重复测量方差分析里的 Mauchly’s test

现在我们来看看 Mauchly’s test 球形假设的检验。根据经验，如果Sig>0.05，那我们数据就有球形度。这里Sig=0,54, 球形没有问题

Mauchly's Test of Sphericity[a]

Measure: rating

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Epsilon[b] Huynh-Feldt	Epsilon[b] Lower-bound
commercial	.898	4.045	5	.543	.942	1.000	.333

“Mauchly's test, $\chi^2(5) = 4.05$, p = 0.54 did **not** indicate any violation of sphericity.”

```
import pingouin as pg
```

Mauchly test for sphericity ¶

“Sphericity is the condition where the variances of the differences between all combinations of related groups (levels) are equal. Violation of sphericity is when the variances of the differences between all combinations of related groups are not equal.” - https://statistics.laerd.com/statistical-guides/sphericity-statistical-guide.php

returns

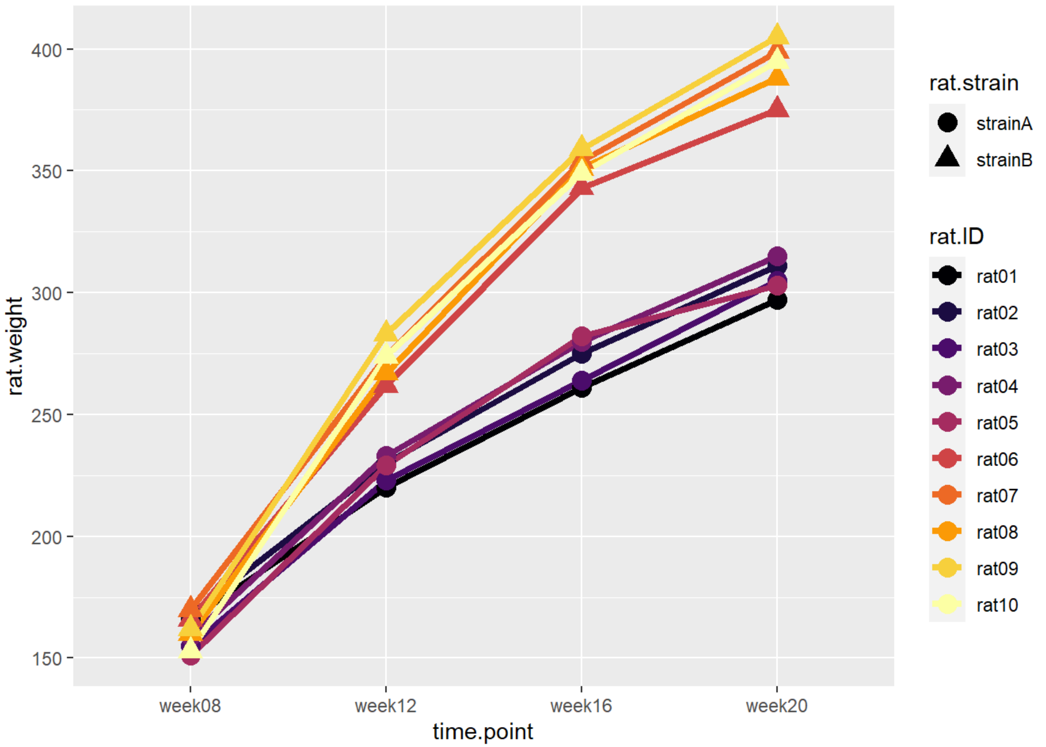
- spher : boolean True if data have the sphericity property.
- W : float Test statistic
- chi_sq : float Chi-square statistic
- ddof : int Degrees of freedom
- p : float P-value.

```
pg.sphericity(data)
```

双因素重复测量方差 two-way repeated measure

双因素重复测量ANOVA（或双因素重复测量ANOVA）与单因素ANOVA非常相似，不同之处在于数据集中的组与重复测量ANOVA中的组之间存在相关性。这意味着受试者在不同的时间、不同的环境或治疗中进行了反复测量。

举一个例子，10只大鼠（**5只来自A菌株，5只来自B菌株**）称重4次，间隔4周（第8周至第20周）



假设是：

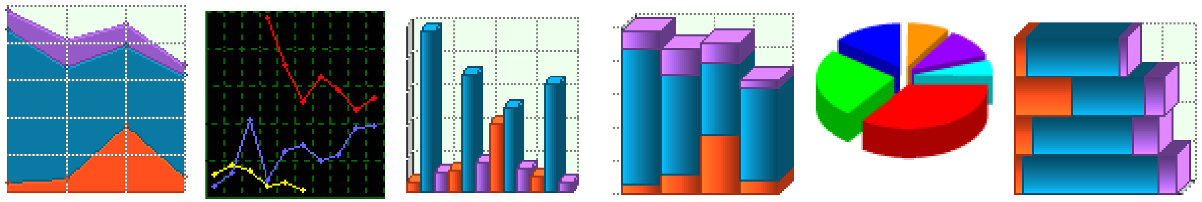
- 每个个体在每个测试组中以测量的形式表示（不存在任何缺失值），
- 分布的正态性（使用Shapiro-Wilk），
- 球形度，这意味着在实验设计中比较任何两组（所有可能的组对）时（通常通过Mauchly试验进行检查），方差齐性验证
- 组不包含异常值

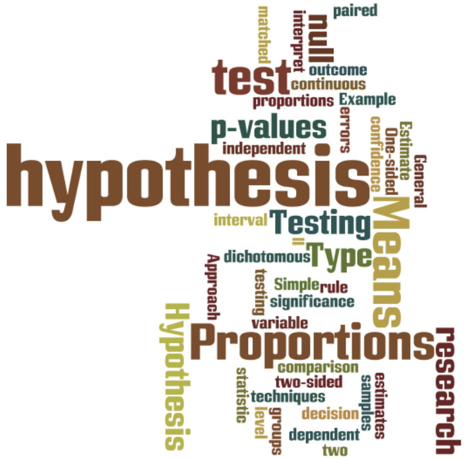
	numDF	denDF	F-value	p-value
(Intercept)	1	24	22147.089	<.0001
time.point	3	24	1750.535	<.0001
rat.strain	1	8	217.521	<.0001
time.point:rat.strain	3	24	94.598	<.0001

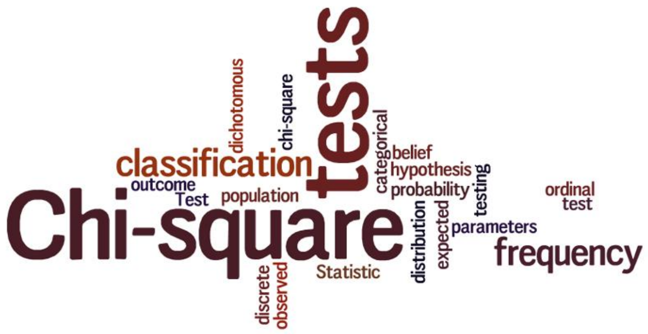
课程内容安排

上课日期	章节	教学内容	教学要点	作业	随堂测	学时
2.16	1	数据可视化，描述性统计	1. 课程介绍 & 数据类型	作业1 (8%)	测试1 (8%)	2
2.23			2. 描述性统计Descriptive Statistics & 数据常用可视化			2
3.2			3. 大数定理 & 中心极限定理			2
3.9			4. 常用概率分布			2
3.16	2	推断性统计，均值差异检验	5. 统计推断基础-1：置信区间 Confidence Interval *	作业2 (10%)	测试2 (10%)	2
3.23			6. 统计推断基础-2：假设检验，I及II类错误，统计量，p-值			2
3.30			7. 数值数据的均值比较-1： 单样本及双样本t-检验，效应量，功效			2
4.6			8. 数值数据的均值比较-2：One-Way ANOVA，正态性检验			2
4.13			9. 数值数据的均值比较-3：Two-Way ANOVA			2
➡ 4.20	3	比例差异检验	10. 样本和置信区间预估 *	作业3 (6%)	测试3 (6%)	2
4.27			11. 类别数据的比例比较-1： 单样本比例推断			2
5.7			12. 类别数据的比例比较-2： 联立表的卡方检验			2
5.11	4	协方差，相关分析，回归分析	13. 相关分析（Pearson r, Spearman rho, Kendal’ s tau) *	作业4 (6%)	测试4 (6%)	2
5.18			14. 简单回归分析			2
5.25			15. 多元回归Multiple Regression			2
6.1	5	Course Summary	16. 课程总结 *			2
			Total	30%	30%	32

*** 随堂测试**



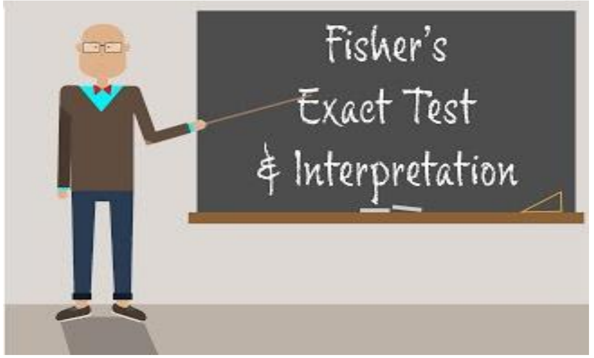




单元 3（week 10, 11, 12）

Inferential Statistics of Proportion Data

- Sample proportion（样本的构成比）
  - Confidence interval estimation for population proportion
  - Confidence interval estimation for difference between 2 populations
- Contingency table（列联表）
- Chi-square test（卡方检验）
  - Chi-Square Test for independence（2 variables, 1 population）
  - Chi-Square Test for goodness of fit（1 variable, 1 population）
- Fisher's exact test（Fisher精确检验）



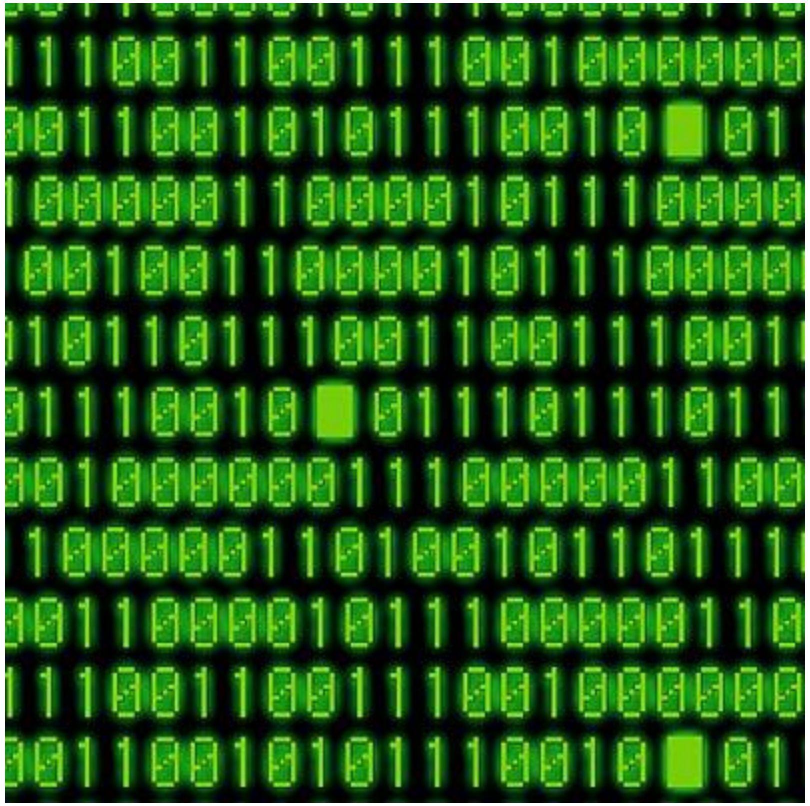
（对应参考书第8章）

Binary Variables 二元变量

Binary variable ≡ two possible outcomes: "success" or "failure"

Examples

- Current smoker (Y/N)
- Gender (M/F)
- Survival 5+ years (Y/N)
- Case or non-case (Y/N)



Binomial Proportions 二项式比例数据

Start by calculating sample proportion “p-hat”:

$$\hat{p} = \frac{x}{n} \quad \text{where } x \equiv \text{no. of successes}$$

Illustration: An SRS of 57 adults identifies 17 smokers. Therefore, the sample proportion is:

$$\hat{p} = \frac{x}{n} = \frac{17}{57} = .2982$$

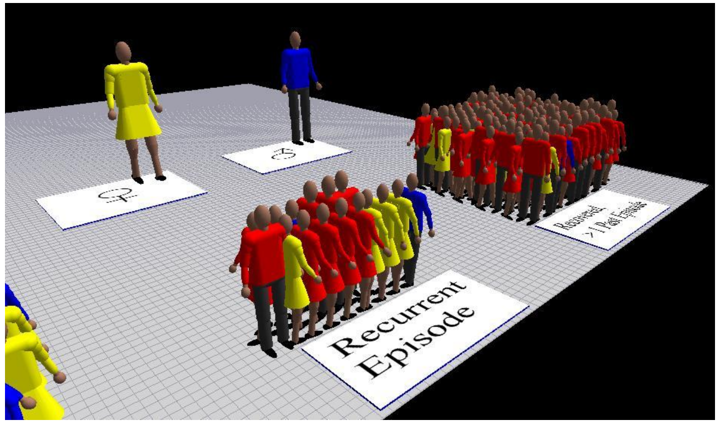
Two Special Proportions

Incidence proportion ≡ proportion at risk that develop a condition over a specified period of time ≡ “average risk”

发病率 ≡ 在特定**时间段**内**从没病到产生疾病**的人的比例 ≡ “平均风险”

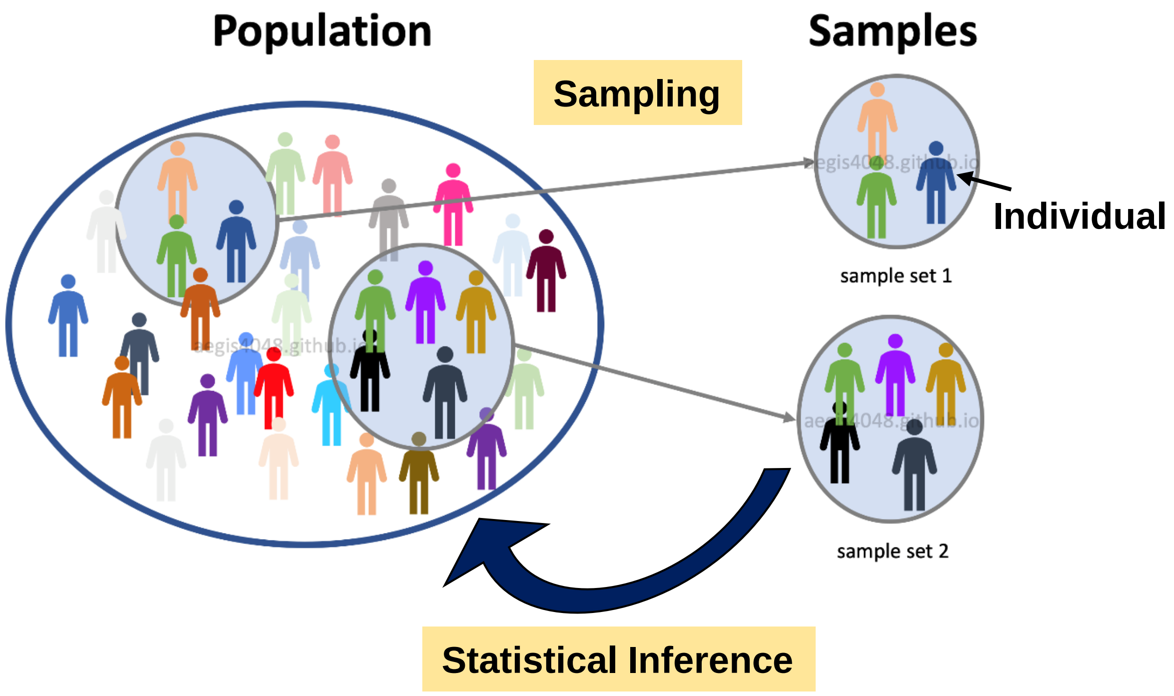
Prevalence proportion ≡ proportion with the characteristic or condition at a particular time

患病率 ≡ 特定时间段里**有某种病**的人的比例



总体和样本

总体：包含指定组的所有成员的数据集。例如：所有居住在中国的人。
样本：包含人口一部分或一部分的数据集。例如：居住在中国的某些人。



我们通过 使用 样本统计量（statistic） 估计 总体参数（parameter） 来进行 统计推断

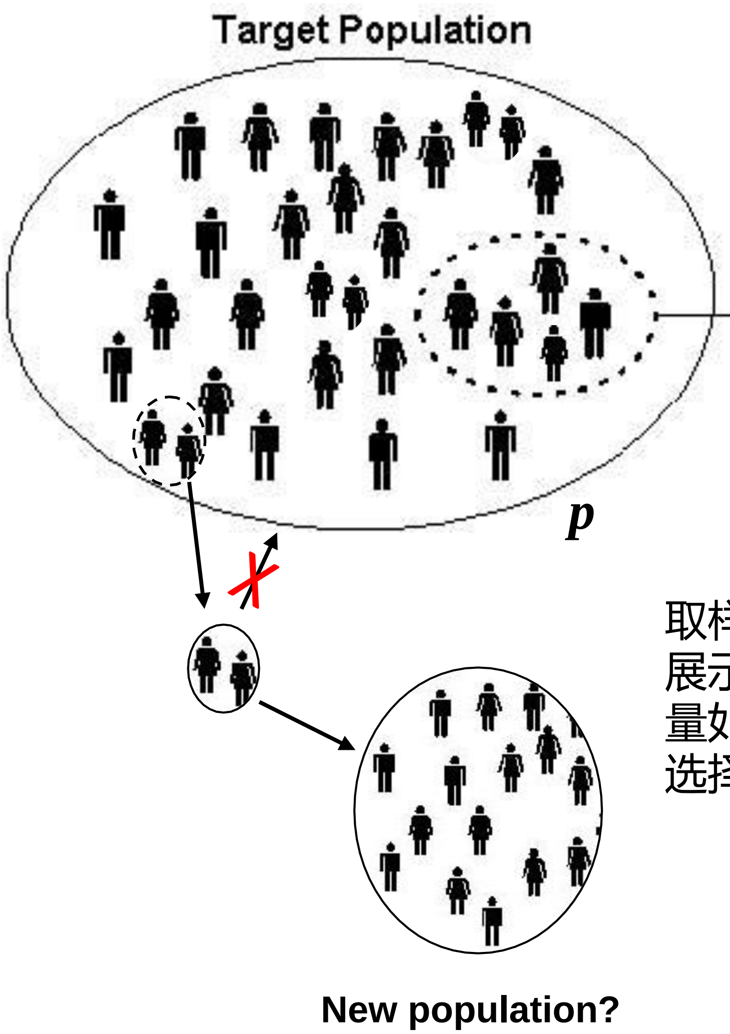
Population versus Sample

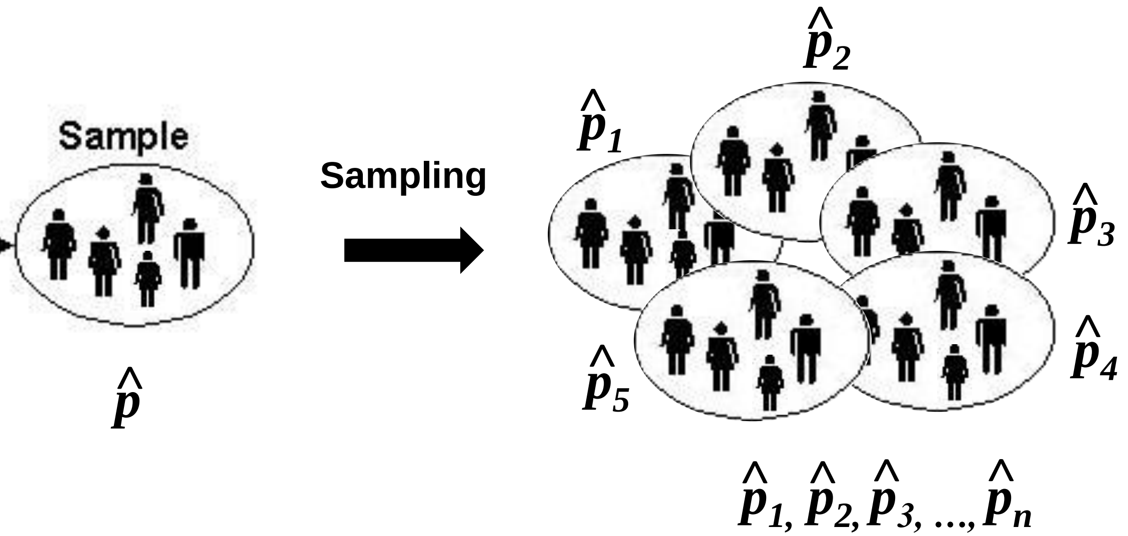
	Parameter		Statistic	
Mean	μ	mu	$\bar{x}$	x-bar
Proportion	p		$\hat{p}$	p-hat
Std. Dev.	σ	sigma	s	
Correlation	ρ	rho	r	

我们使用样本统计量进行统计推断的方法称为估算/估计（estimation）

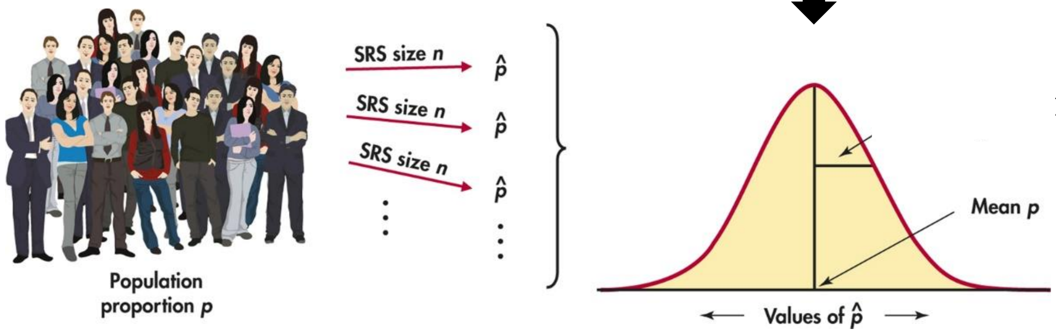
取样分布（Sampling distribution）

	Parameter		Statistic	
Mean	μ	mu	$\bar{x}$	x-bar
Proportion	p		$\hat{p}$	p-hat
Std. Dev.	σ	sigma	s	
Correlation	ρ	rho	r	



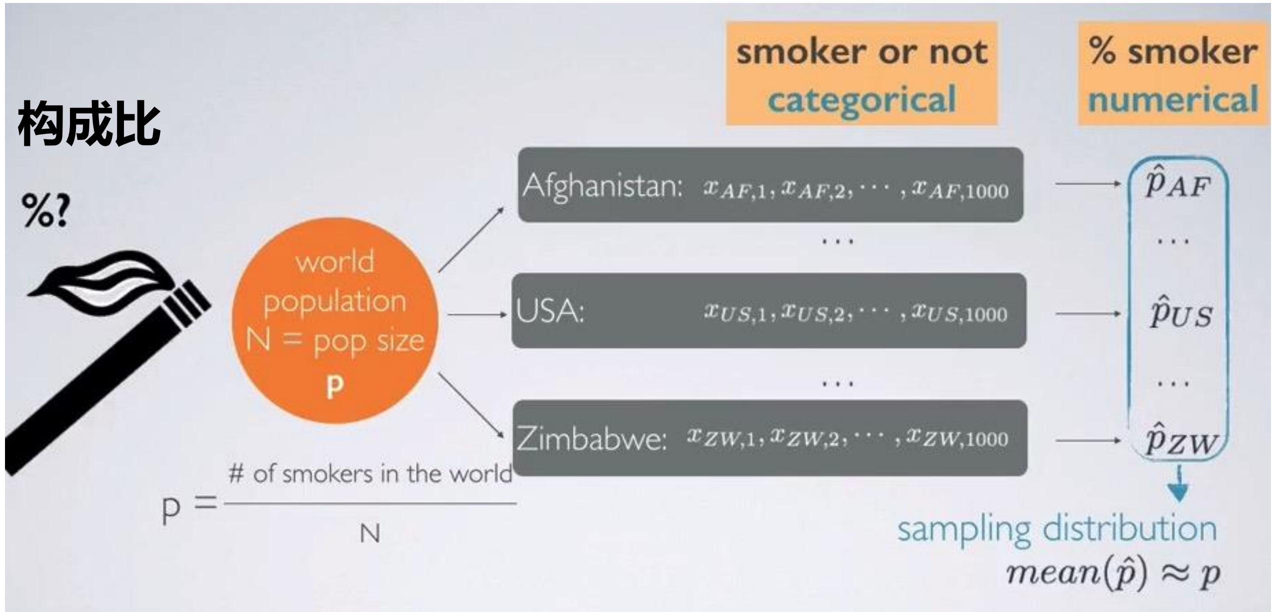


取样分布向我们展示了样本统计量如何随样本的选择而变化



不管总体的分布是怎样的，二项分布，伯努利分布，指数分布，还是什么。比列值取样分布将近似于一个正态分布
(CLT)

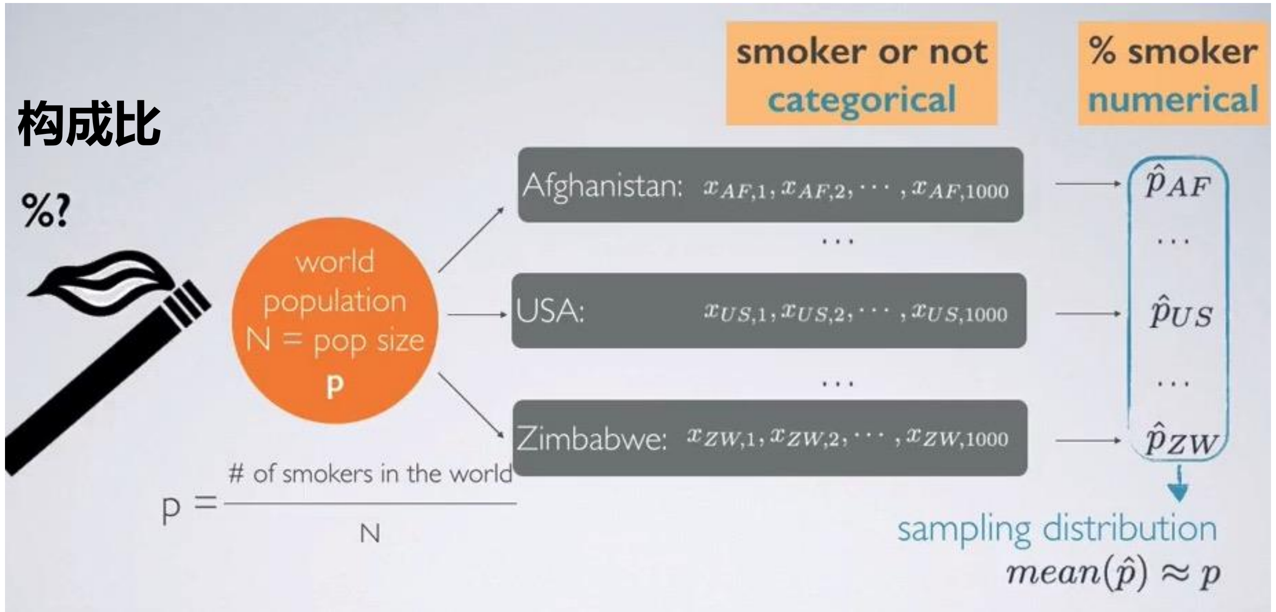
Sampling Distribution of a Sample Proportion 构成比的取样分布



分类变量为：吸烟者与非吸烟者

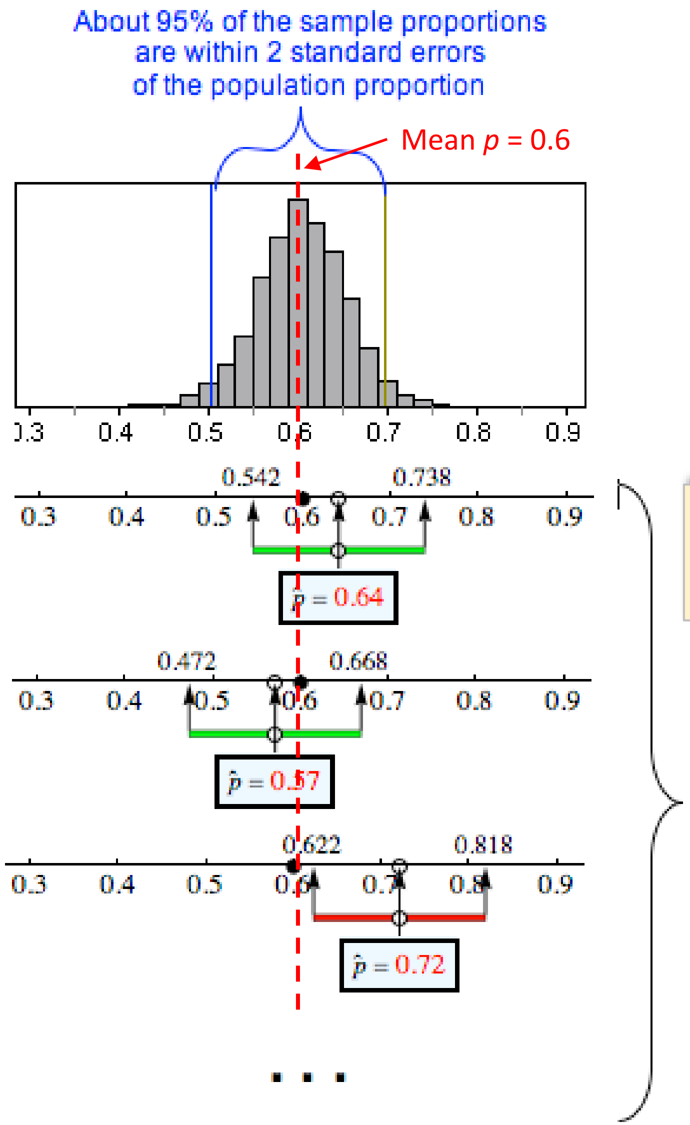
- 因不知道人口比例和规模，因此对每个国家都做了估计。从每个国家抽取1000个样本并计算出比例值
- 多个比例将会构成抽样分布，于是比例的均值将是总体人口的近似比例
- 因此，一开始是1个分类变量，但可以转换它的样本统计为数值变量

Sampling Distribution of a Sample Proportion 构成比的取样分布



分类变量为：吸烟者与非吸烟者

- 因不知道人口比例和规模，因此对每个国家都做了估计。从每个国家抽取1000个样本并计算出比例值
- 多个比例将会构成抽样分布，于是比例的均值将是总体人口的近似比例
- 因此，一开始是1个分类变量，但可以转换它的样本统计为数值变量



根据采集一个样本，然后我们想获得整体抽烟率的置信区间CI

When margin of error is 2 standard errors, about 95% of confidence intervals contain the population proportion.

晚期黑色素瘤

晚期黑色素瘤是一种侵袭性皮肤癌，直到最近几乎都是致命的。

目前正在研究能触发针对癌症的免疫反应、从而导致黑色素瘤停止进展或完全消失的疗法(CAR-T）。

在一项研究中，52名患者同时接受了两种新疗法，即nivolumab和ipilimumab，其中21名患者出现了免疫反应。（Wolchok, et. al. *NEJM* (2013) 369(2): 122-33.）

通过推论性统计可以问以下几个常见问题 …

- What is the estimated population probability of immune response following concurrent therapy with nivolumab and ipilimumab? 同时接受nivolumab和ipilimumab治疗后，**估计**在人群中产生免疫反应的概率是多少?
- What is the 95% confidence interval for the estimated population probability of immune response following concurrent therapy with nivolumab and ipilimumab? nivolumab和ipilimumab同时治疗后，**估计**人群中免疫反应概率的95%置信区间是多少?
- In previous studies, the proportion of patients responding to one of these agents was 30% or less. Do these results suggest that the probability of response to concurrent therapy is better than 0.30? 在前者的研究中，对任何其中一种药物有反应的患者比例不超过30%。现在这些结果是否表明对同时治疗的反应概率优于0.30?

二项式比例数据的推论统计（Inference for binomial proportions）

前面讲的黑色素瘤治疗数据是个二项数据，其“成功 success”为“产生了免疫反应”

假设X是一个参数为n和p的二项式随机变量，其中n就是试验次数，p是成功概率

- 这里的推断统计计算是将针对总体参数 p 进行推断, 即总体中里成功概率
- 根据样本对 p 进行估计 $\hat{p} = x/n$, 其中x是观察到的成功次数

p 的推断统计可以使用二项式的正态近似，或直接使用二项式分布

二项分布（Binomial distribution）

Let x = number of successes in n trials

$$P(x \text{ successes}) = \binom{\#\text{ of trials}}{\#\text{ of successes}} p^{\#\text{ of successes}} (1-p)^{\#\text{ of trials - }\#\text{ of successes}}$$

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, 2, \ldots, n$$

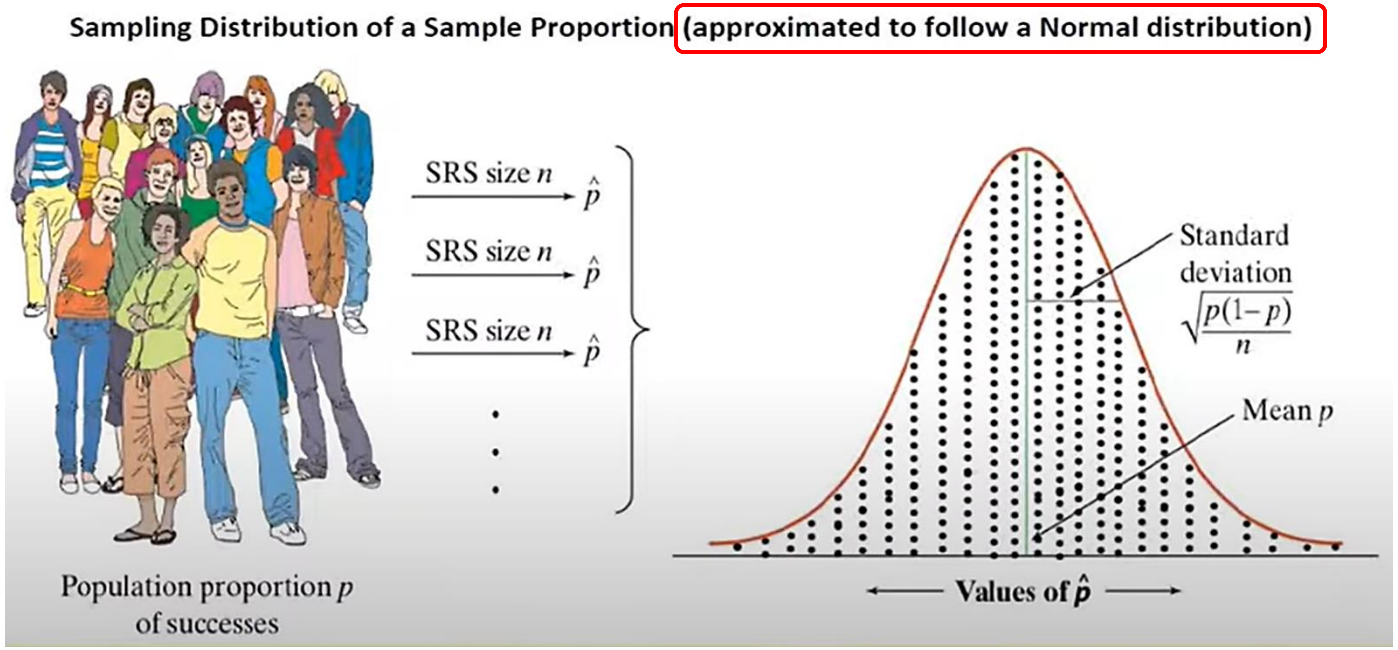
Parameters of the distribution:

- n = number of trials
- p = probability of success

Shorthand notation: $X \sim \text{Bin}(n, p)$

- 二项分布描述了在n个独立的伯努利试验中获得k次成功的概率，成功概率为p。
- 前面课程里举的米尔格拉姆实验例子里，就是计算每4次试验有一次反对的概率是多少。

构成比的取样分布 - 二项式分布



p 的分布与二项分布（Binomial）密切相关。二项分布就是1个总成功数的分布，而 p 分布是**平均成功数**的分布。

The mean of the binomial distribution which we know to be *np*:

$$E(p') = E\left(\frac{x}{n}\right) = \left(\frac{1}{n}\right)E(x) = \left(\frac{1}{n}\right)np = p$$

$$\sigma_{p'}{}^2 = \mathrm{Var}(p') = \mathrm{Var}\left(\frac{x}{n}\right) = \frac{1}{n^2}(\mathrm{Var}(x)) = \frac{1}{n^2}(np(1-p)) = \frac{p(1-p)}{n}$$

Parameter	Population distribution	Sample	Sampling distribution of p's
Mean	$\mu = np$	$p' = \frac{x}{n}$	p' and $E(p') = p$
Standard Deviation	$\sigma = \sqrt{npq}$		$\sigma_{p'} = \sqrt{\frac{p(1-p)}{n}}$

如何从单个比例推断总体比例的置信区间

如果样本的比例为：$\hat{p}$，比如总统选举的民意调查，从N个问卷中，有m个支持票，

那么样本支持率为：

$$\hat{p} = \frac{m}{N}$$

在无法获得全国数据的情况下，如何从$\hat{p}$来估算全国的支持率p的置信区间？

如果把p看作随机变量，根据CLT，在大量重复抽样（每次仍然采样N个）

$$var(p) = var(\frac{m}{N})$$

$$= \frac{1}{N^2} var(m)$$

$\because$ 显然在随机采样的情况，这个采样过程符合二项式分布$X \sim B(N,p)$

$$\therefore var(m) = Np(1-p)$$

$$\therefore var(p) = \frac{1}{N^2} \cdot Np(1-p)$$

$$= \frac{p(1-p)}{N}$$

$$\therefore \sigma_p = \sqrt{\frac{p(1-p)}{N}}$$

σ_p是每次采样N个观察值，支持率p分布$(N(p, \sigma_p^2))$的标准差。

实际中因为p未知，但当总体足够大$(e.g.,N$小于5%总体样本数，$N\hat{p} \geq 10, and\ N(1-\hat{p}) \geq 10)$

我们可以用样本标准差$(\sigma_{\hat{p}})$替代样本比例$\hat{p}$的标准误σ_p：

$$\therefore CI = \hat{p} \pm Z_{\alpha/2} \cdot \sigma_{\hat{p}} = \hat{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$$

Formula

$$CI = \bar{x} \pm z\frac{s}{\sqrt{n}}$$

CI = confidence interval

$\bar{x}$ = sample mean

z = confidence level value

s = sample standard deviation

n = sample size

Confidence intervals for population proportions

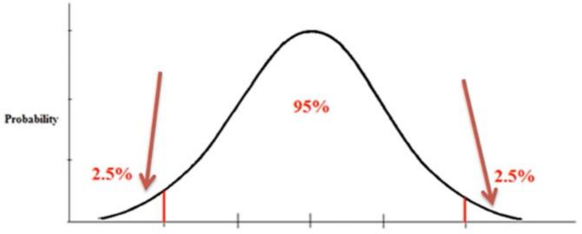
Confidence interval for Population proportion

- Each element in the population can be classified as a success or failure

  Sample proportion $\hat{p} = \dfrac{\text{number of successes}}{\text{sample size}} = \dfrac{x}{n}$
- Proportion always between 0 and 1
- For large samples the sample proportion $\hat{p}$ is approximately normal

$$CI(p)_{1-\alpha} = \left[\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

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



p

$\alpha = 0.05$

$z_{0.05/2} = z_{0.025}$

CI For	Sample Statistic	Margin of Error	Use When
Population mean (μ)	$\bar{x}$	$\pm z^* \dfrac{\sigma}{\sqrt{n}}$	X is normal, or $n \geq 30$; σ known
Population mean (μ)	$\bar{x}$	$\pm t^*_{n-1} \dfrac{s}{\sqrt{n}}$	$n < 30$, and/or σ unknown
Population proportion (p)	$\hat{p}$	$\pm z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$	$n\hat{p},\ n(1-\hat{p}) \geq 10$
Difference of two population means ($\mu_1 - \mu_2$)	$\bar{x}_1 - \bar{x}_2$	$\pm z^* \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$	Both normal distributions or $n_1, n_2 \geq 30$; σ_1, σ_2 known
Difference of two population means $\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$\pm t^*_{n_1+n_2-2} \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$	$n_1, n_2 < 30$; and/or $\sigma_1 = \sigma_2$ unknown
Difference of two proportions ($p_1 - p_2$)	$\hat{p}_1 - \hat{p}_2$	$\pm z^* \sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$n\hat{p}, n(1-\hat{p}) \geq 10$ for each group

Hypothesis test for sample proportion（z-test）

$\hat{p}$ vs. p

	confidence interval	hypothesis test
uccess-failure condition	$n\hat{p} \geq 10$ $n(1-\hat{p}) \geq 10$	$np \geq 10$ $n(1-p) \geq 10$
standard error	$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$SE = \sqrt{\frac{p(1-p)}{n}}$

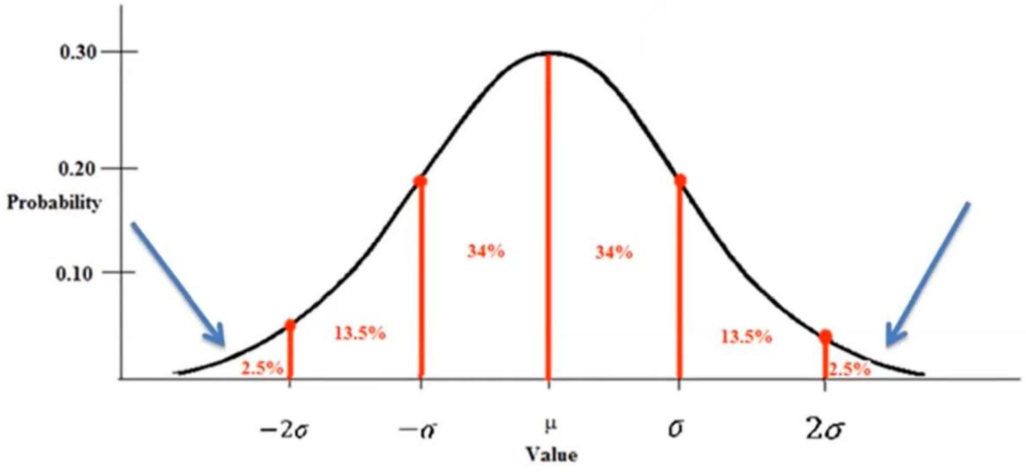
Hypothesis testing for a single proportion:

1. Set the hypotheses: $H_0 : p = null\ value$
   $H_A : p < \ or\ > \ or\ \neq\ null\ value$
2. Calculate the point estimate: $\hat{p}$
3. Check conditions: 当以下条件成立时，p^的抽样分布近似为正态分布：
   1. ***Independence:*** Sampled observations must be independent (random sample/assignment & if sampling without replacement, n < 10% of population)
   2. ***Sample size/skew:*** $np \geq 10$ and $n(1-p) \geq 10$ 样本中预计至少有10次成功和10次失败
4. Draw sampling distribution, shade p-value, calculate test statistic $Z = \frac{\hat{p}-p}{SE}$, $SE = \sqrt{\frac{p(1-p)}{n}}$
5. Make a decision, and interpret it in context of the research question:
   - If p-value < α, reject H_0; the data provide convincing evidence for H_A.
   - If p-value > α, fail to reject H_0 the data *do not* provide convincing evidence for H_A.

如果H0被拒绝，只能说本样本不足以支持H0成立
如果H0不被拒绝, 也并不能说明H0一定正确，只能说本数据不足以反驳H0

Hypothesis test for sample proportion (z-test)

➢ One Sample z-Test for Proportions



A survey claims that 9 out of 10 doctors recommend aspirin for their patients with headaches. To test this claim, a random sample of 100 doctors is obtained. Of these 100 doctors, 82 indicate that they recommend aspirin. Is this claim accurate? Use alpha = 0.05

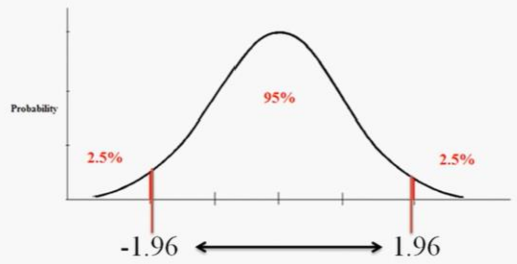
1. **Define Null and Alternative Hypotheses**

$$H_0; p = .90$$
$$H_1; p \neq .90$$

2. **State Alpha**

$$\alpha = 0.05$$

3. **State Decision Rule**



If z is less than -1.96 or greater than 1.96, reject the null hypothesis.

4. **Calculate Test Statistic**

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$\hat{p} = .82$$
$$p_0 = .90$$
$$n = 100$$

$$z_0 = \frac{.82 - .90}{\sqrt{\frac{.90(1-.90)}{100}}} = \frac{-0.08}{0.03} = -2.667$$

5. **State Results**

Decision Rule: If Z is less than -1.96, or greater than 1.96, reject the null hypothesis.

Z = -2.667

Result: Reject H0

6. **State Conclusion**

The claim that 9 out of 10 doctors recommend aspirin for their patients is not accurate, z = -2.667, p < 0.05.

Hypothesis test for proportions, two samples

Conditions for inference for comparing two independent proportions:

1. ***Independence:***
   - ✓ **within groups:** sampled observations must be independent within each group
     - random sample/assignment
     - if sampling without replacement, n < 10% of population
   - ✓ **between groups:** the two groups must be independent of each other (non-paired)
2. ***Sample size/skew:*** Each sample should meet the success-failure condition:
   - ✓ $n_1 p_1 \geq 10$ and $n_1(1-p_1) \geq 10$
   - ✓ $n_2 p_2 \geq 10$ and $n_2(1-p_2) \geq 10$

estimating the difference between two proportions

point estimate ± margin of error

$$(\hat{p}_1 - \hat{p}_2) \pm z^{\star} SE_{(\hat{p}_1 - \hat{p}_2)}$$

Standard error for difference between two proportions, for calculating a confidence interval:

$$SE = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

➢ z-Test for Proportions, Two Samples

Researchers want to test the effectiveness of a new anti-anxiety medication. In clinical testing, 64 out of 200 people taking the medication report symptoms of anxiety. Of the people receiving a placebo, 92 out of 200 report symptoms of anxiety. Is the medication working any differently than the placebo? Test this claim using alpha = 0.05

1. Define Null and Alternative Hypotheses

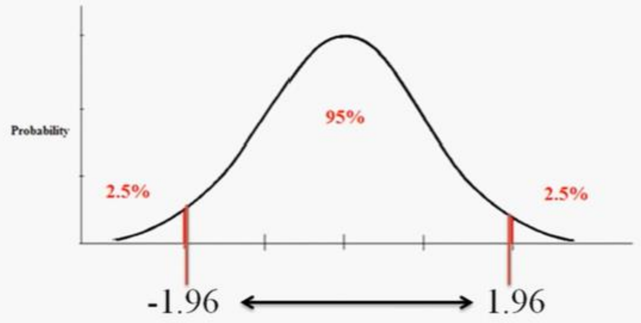
$$H_0; p_1 = p_2$$

$$H_1; p_1 \neq p_2$$

2. State Alpha

α = 0.05

3. State Decision Rule



If z is less than -1.96 or greater than 1.96, reject the null hypothesis.

4. Calculate Test Statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$n_1 = 200$ $\quad \hat{p}_1 = \frac{64}{200} = 0.32$

$n_2 = 200$ $\quad \hat{p}_2 = \frac{92}{200} = 0.46$

$\hat{p} = 0.39$

$$z = \frac{(0.32 - 0.46)}{\sqrt{0.39(1-0.39)}\sqrt{\frac{1}{200} + \frac{1}{200}}} = \frac{-0.14}{(.488)(0.1)} = 2.869$$

5. State Results

Decision Rule: If z is less than -1.96, or greater than 1.96, reject the null hypothesis.

z = 2.869

Result: Reject H0.

6. State Conclusion

There was a significant difference in effectiveness between the medication group and the placebo group, z = -2.869, p < 0.05.

谢谢，下周见！