# observing behaviour

31 October 2021 14:29

## Alphabet

| | | |
|---|---|---|
| Pooled variance estimate | 联合方差 | 两组sample，size不同的时候用。size相同则不需要这个步骤 |
| Weighted average | 加权平均 | |

## Content

1.

# Summary

04 December 2021 16:08

## Confidence/significance level/P value

1. https://baike.baidu.com/item/%E6%98%BE%E8%91%97%E6%80%A7%E6%B0%B4%E5%B9%B3/1383148?fr=aladdin
2. https://baike.baidu.com/item/%E7%BD%AE%E4%BF%A1%E6%B0%B4%E5%B9%B3/7442548?fr=aladdin
3. Level of confidence: 置信水平, β, 有总体特征的事件发生占总数的概率, 较大的值(0.95)
4. Level of significance: 显著水平, α, 小概率事件发生的概率, 较小的值(0.05)
5. 显著水平和置信水平之和为1
6. P: 发生一类错误的概率, 小于显著水平α就可以拒绝NH

## 单侧与双侧的选取

1. 楼上已经解释了单双侧检验的区别。但是貌似没有说一般什么时候用单侧什么时候用双侧。

   例如：在计量经济学当中，你的假设如果是 $H_0 : \hat{\beta}_1 = k, H_1 : \hat{\beta}_1 \neq k.$ 这种情况一般用双侧（k是一个固定值）

   如果是 $H_0 : \hat{\beta}_1 \leq k, H_1 : \hat{\beta}_1 \geq k.$ 这种一般是单边检验
2. 比大小用单侧, 判断是否相等用双侧
3. 双侧检验只能证明你比较的两个项目之间是有差异的，但是不能告诉你这两组中哪个高哪个低；单侧检验就可以告诉你是低于还是高于。

## 几种检验的比较

1. Reference
   a. https://blog.csdn.net/tianguiyuyu/article/details/80789856
   b. https://zhuanlan.zhihu.com/p/49468324
2. 单尾/双尾
   a. https://keydifferences.com/difference-between-one-tailed-and-two-tailed-test.html#KeyDifferences

b.

| BASIS OF COMPARISON | ONE-TAILED TEST | TWO-TAILED TEST |
|---|---|---|
| Meaning | A statistical hypothesis test in which alternative hypothesis has only one end, is known as one tailed test. | A significance test in which alternative hypothesis has two ends, is called two-tailed test. |
| Hypothesis | Directional | Non-directional |
| Region of rejection | Either left or right | Both left and right |
| Determines | If there is a relationship between variables in single direction. | If there is a relationship between variables in either direction. |
| Result | Greater or less than certain value. | Greater or less than certain range of values. |
| Sign in alternative hypothesis | > or < | ≠ |

3. T/chi
   a. https://www.reference.com/world-view/difference-between-t-test-chi-squared-test-5816dd2e022c8
   b.
      1. T-test allows you to differentiate between the two groups. While the Chi-square test also helps you to find the relationship between two variables but has no direction and size of the relationship.
      2. Null hypothesis: In the T-test, there is no stat. difference between the two groups while in the Chi-square test there is no relationship between two variables.
   c. NH是正好相反的, 卡方检验是假设无相关性, t检验是假设无差别(null difference, 从这里衍生出了NH)
4. 各种t
   a. One-sample
      i. A student can use a one-sample t-test or a two-sample t-test. A one-sample t-test is designed to answer a null hypothesis that concerns the data set's mean when the data are from independent observation and follow a normal distribution.
   b. Two-sample
      i. The two-sample t-test evaluates the null hypothesis when two sets of data are collected. Both sets of data must come from the same sample size for the results to be valid.
5. 各种chi
   a. https://www.codementor.io/@abhirajsuresh/what-is-a-chi-square-test-and-why-do-we-

use-it-1365snyolr

b. Chi-Square Goodness of Fit Test

   i. This is a non-parametric test. We typically use it to find how the observed value of a given event is significantly different from the expected value. In this case, we have categorical data for one independent variable, and we want to check whether the distribution of the data is similar or different from that of the expected distribution.

c. Chi-Square Test for Association/Independence

   i. The second type of chi-square test is the **Pearson's chi-square test** of association. This test is used when we have categorical data for two independent variables and we want to see if there is any relationship between the variables.

d.
- The Chi Square test is useful when the observations take the form of counts (how many times an event of interest occurs);
- One way classification can show how well the observations fit the expectations;
- Two way classification can show how much two variables of interest are associated.

6. Z检验

   a. https://baike.baidu.com/item/z%E6%A3%80%E9%AA%8C/8374223?fr=aladdin

   b. 用法1: 总体标准差已知或样本容量大于30，比较两个样本的均值是否有显著性的差异

      i.
$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(\bar{X}_1-\bar{X}_2)-0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}=\frac{(\bar{X}_1-\bar{X}_2)-0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

$\bar{X}_1/\bar{X}_2$为两个样本的均值；

$\mu_1/\mu_2$为两个样本的抽样总体的均值，检验时假设两个总体的均值相等，所以差为0；

$\sigma_1^2/\sigma_2^2$为两个总体的标准差；

$S_1^2/S_2^2$为两个样本的标准差；



   c. 用法2: 总体标准差已知或样本容量大于30，比较某个总体的均值与某个常数是否有显著性的差异

      i.
$$Z=\frac{\bar{X}_i-\mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z=\frac{\bar{X}_i-\mu}{\frac{S}{\sqrt{n}}}$$

$\bar{X}_i$为样本的均值；

$\mu$为假设与样本均值无显著性差异的常数；

$\sigma$为总体标准差；

$S$为样本标准差；



7. Z与卡方

# Lecture 1: introduction

05 October 2021 23:41

- The course involves an Assessed Exercise that accounts for 20% of the final mark;
- The course includes 10 hours (out of 30) that will be dedicated to the Assessed Exercise;
- The exam accounts for 80% of the final mark.

## The theory of multiple intelligence

1. Individuals possess eight or more relatively autonomous intelligences and draw on these intelligence to solve problems that are relevant to the societies in which they live
    a. Isolation
    b. Specific neural structure
    c. Distinct development
    d. Evolutionary basis
    e. Symbolic systems
    f. Measurability
    g. Experimental
    h. Specific processes

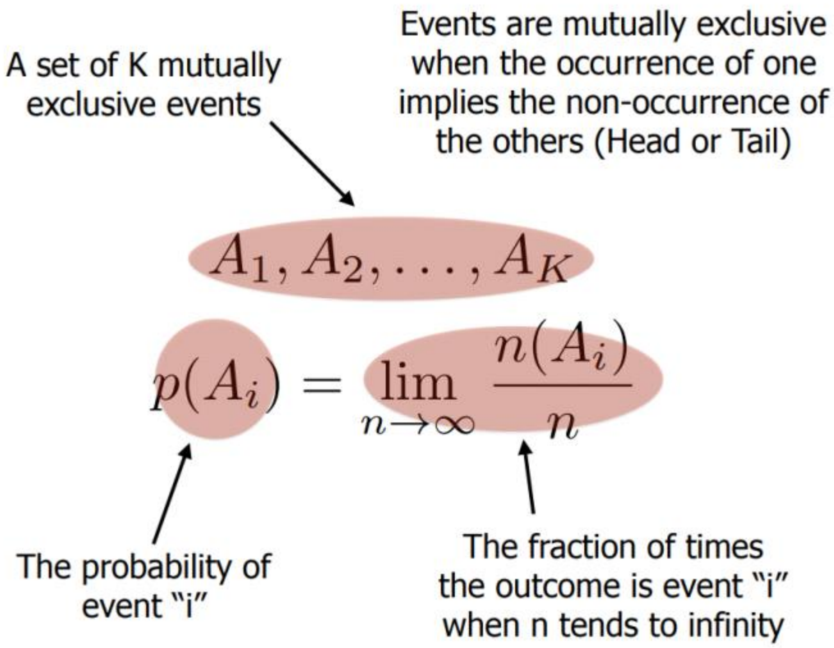
## Type of intelligence

1. Linguistic
2. Logical-mathematical
3. Spatial
4. Musical
5. Naturalist
6. Body-kinesthetic
7. Interpersonal
8. Intrapersonal

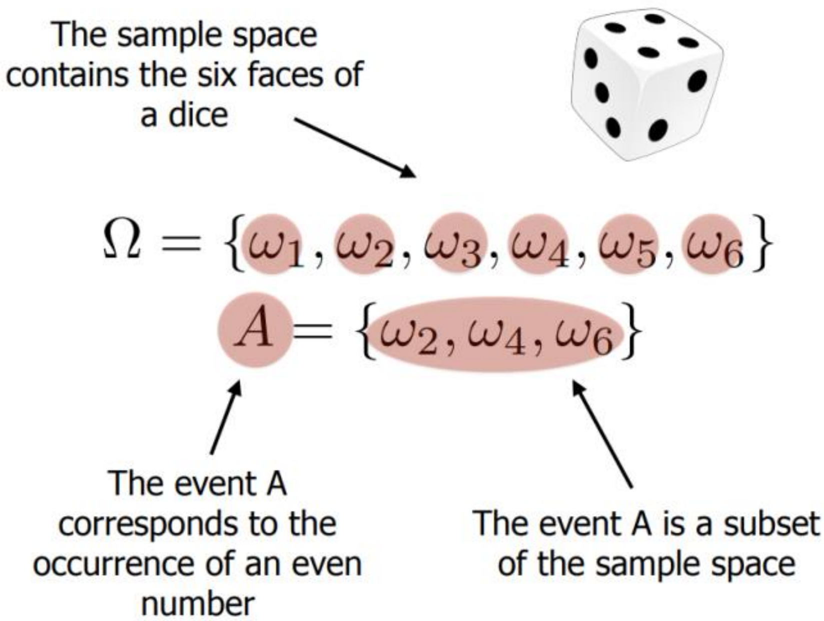
# Lecture 2: basis statistics

05 October 2021 23:56

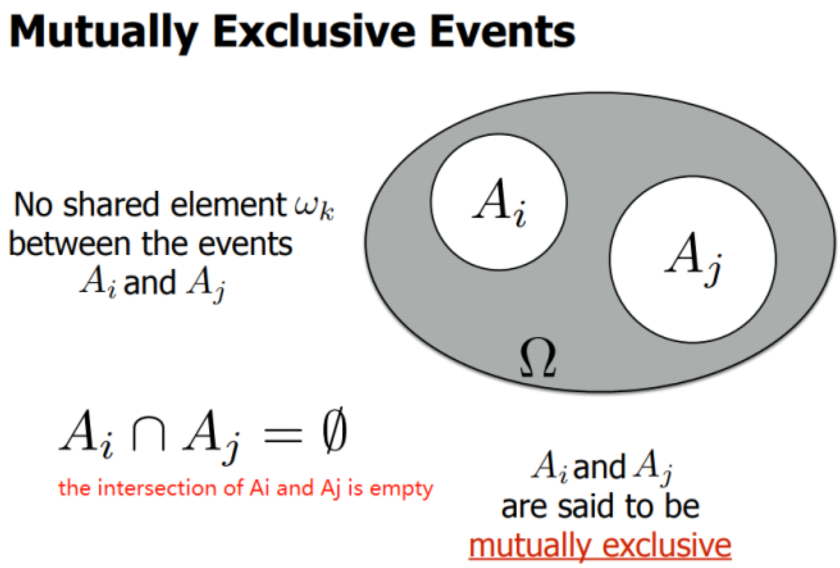
## Definition of probability

1. 

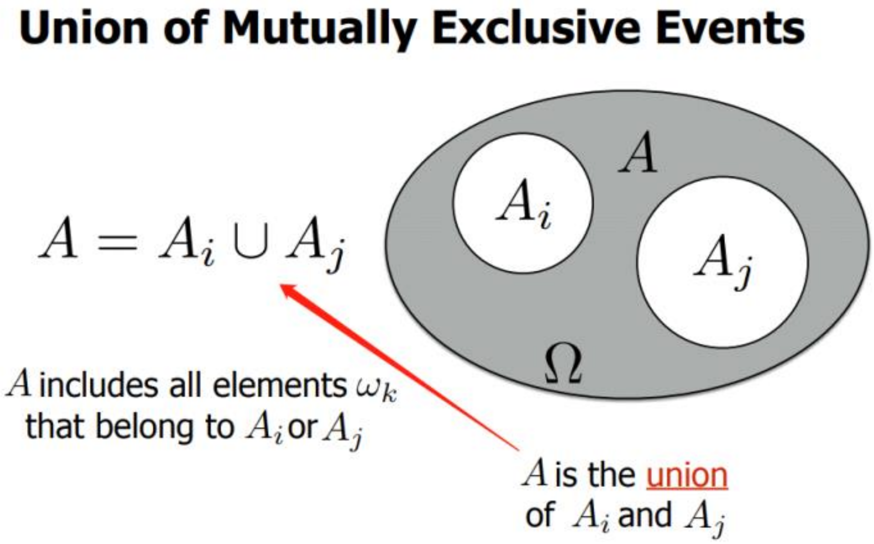

## Basic laws of probability

1. The sample space
   a. 

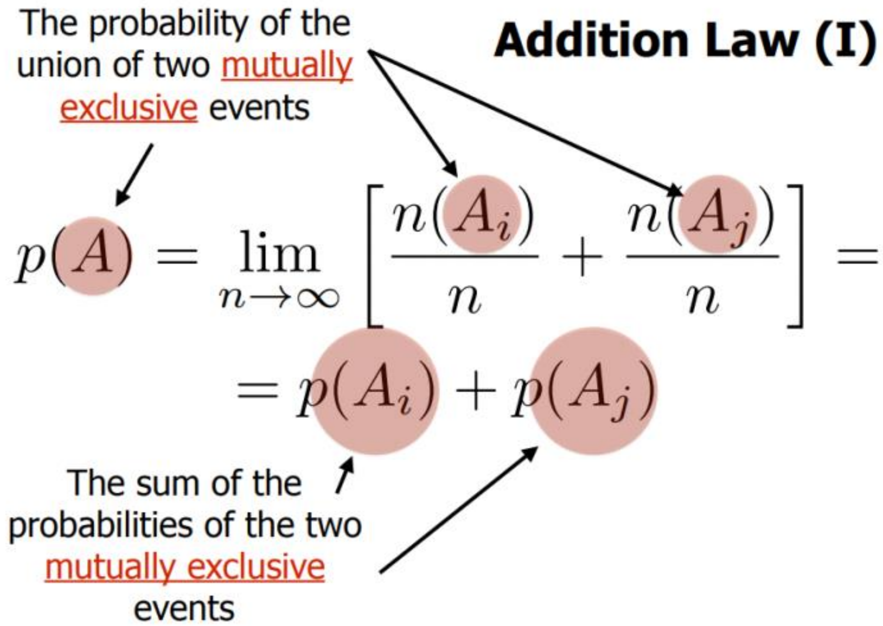
2. Mutually exclusive events(互斥事件)

a. 


3. Union of mutually exclusive events

a. 


4. Additional law 1

a. 


b. 互斥事件发生的概率等于他们子事件发生概率之和

5. Additional law 2

## Addition Law (II)

A is the union of N mutually exclusive events

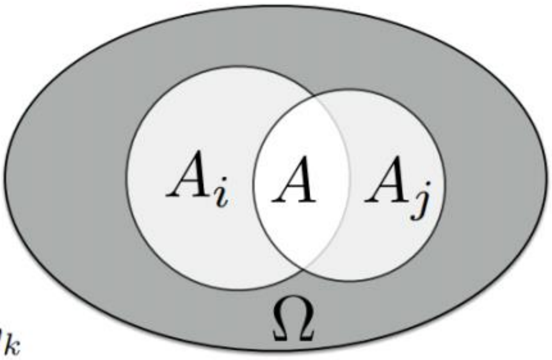
a. $$p(A) = p\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} p(A_n)$$

p(A) is the sum of the probabilities of the mutually exclusive events

6. Intersection of multiple events

## Intersection of Multiple Events

a. $$A = A_i \cap A_j$$




$A$ includes all elements $\omega_k$ that belong to both $A_i$ and $A_j$

$A$ is the intersection of $A_i$ and $A_j$

7. Conditional probability(条件概率)

The probability of A given B

Probability of the intersection of A and B

a. $$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Probability of B

b. Product law

c.

**Product Law**

The probability of the intersection

$$p(A \cap B) = p(A, B) = p(A|B)p(B)$$

The joint probability of A and B

The product comes from the definition of the conditional probability

d.

The intersection is a subset of the two sets

$$A \cap B \subseteq B \Rightarrow 0 \leq p(A|B) \leq 1$$

$$B \subset A \Rightarrow p(A|B) = 1$$

$$A \cap B = \emptyset \Rightarrow p(A|B) = 0$$

The intersection is empty

e.

$$p(A|B) = \sum_{n=1}^{N} \frac{p(A_n \cap B)}{p(B)} =$$

$$= \sum_{n=1}^{N} p(A_n|B)$$

f. $$p(A) = \sum_{n=1}^{N} p(A \cap B_n) = \\ = \sum_{n=1}^{N} p(A|B_n)p(B_n)$$

g.

The definition of probability

$$p(A_1, A_2) = \lim_{n\to\infty} \frac{n(A_1, A_2)}{n} \simeq \\ \simeq \lim_{n\to\infty} \frac{n(A_1, A_2)}{n(A_2)} \simeq p(A_1)$$

The number of times A2 occurs tends to infinite when n does

If the occurrence of A2 does not influence the occurrence of A1

h.

$$p(A_1, A_2) = \lim_{n\to\infty} \frac{n(A_1, A_2)}{n} \simeq \\ \simeq \lim_{n\to\infty} \frac{n(A_1, A_2)}{n(A_2)} \frac{n(A_2)}{n} \simeq p(A_1)p(A_2)$$

Multiply and divide by n(A2)

Only if the occurrence of A2 does not influence the occurrence of A1 and vice versa

i.

The sum goes over all values of the discrete random variable

This is a constraint that must be respect for p to be a distribution

$$\sum_{k=1}^{T} p(x_k) = 1$$

j.

The probability of a variable being lower than a certain value

The Addition Law allows the following estimate

k. $$P(\xi \leq x_j) = \sum_{x_k \leq x_j} p(x_k)$$

$$F(x_j) = P(\xi \leq x_j)$$

The cumulative probability function

The value of the cumulative probability function is a probability

A variable is continuous when its value belongs to a continuous interval

$$\xi \in [a, b]$$

l. $$P(x_1 \leq \xi \leq x_2) = \int_{x_1}^{x_2} p(x)dx$$

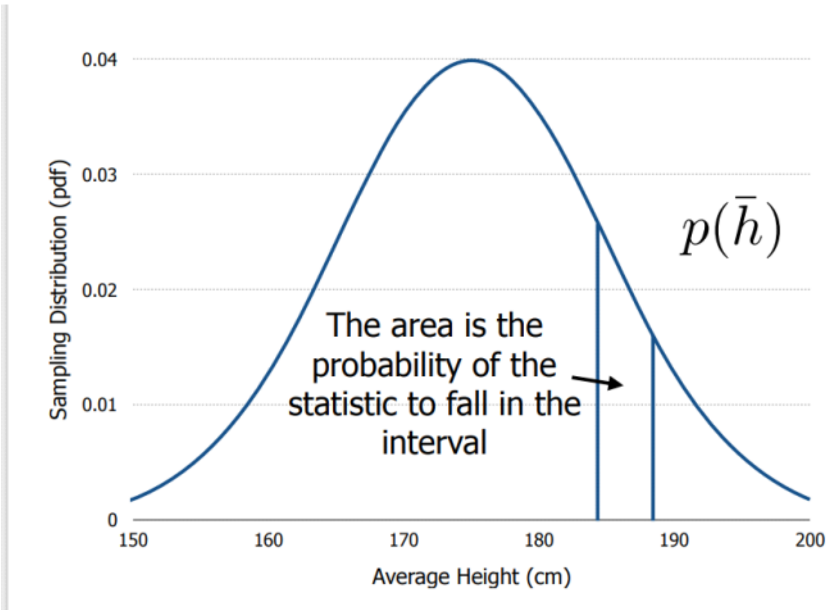
It is possible to estimate the probability of the value falling between two extremes

The function p(x) is called probability density function

# Lecture 3: hypothesis testing

21 October 2021 17:46

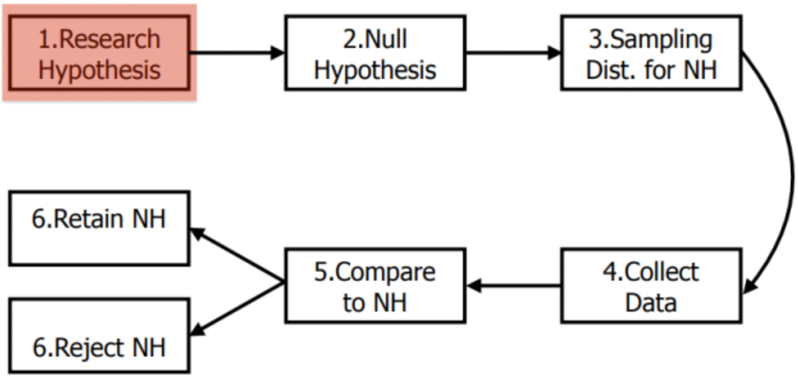
## Sampling Distribution

1. Reference
   a. https://zhuanlan.zhihu.com/p/36426590
   b. https://zhuanlan.zhihu.com/p/216410693
2. Statistical testing：针对人行为的调查
   a. Statistical testing is a process aimed at comparing different populations in terms of statistical properties;
3. "[...] the distribution of values obtained for [a] $\boxed{\text{statistic over repeated sampling}}$
4. 


## Hypothesis Testing

1. Reference
   a. https://zhuanlan.zhihu.com/p/86178674
   b. https://wiki.mbalib.com/wiki/%E5%81%87%E8%AE%BE%E6%A3%80%E9%AA%8C
2. Workflow

a.
- State a research hypothesis;
- Setup the null hypothesis (NH);
- Construct the sampling distribution of the statistic when the null hypothesis is true;
- Collect data;
- Compare sample statistic to its distribution when the null hypothesis is true;
- Retain or reject the null hypothesis.

b. 


3. Confidence level(置信区间)
   a. 显著水平（critical value）0.05或0.01
   b. 落在最高或者最低的0.05内为不可靠数据，不能排除NH
   c. 在这个区间之内的是可靠数据，可以排除NH
   d. 如果担心损失的数据太多了，如选取0.05的话就相当于排除了10%的数据，可以两头都取0.025，加在一起够就好
   e. 只取一边的为one-tailed directional test, 两边都取的two-tailed nodirectional test
   f. 通常选择two-tailed test
4. Type 1 error & type 2 error
   a. Reference
      i. https://www.cnblogs.com/leezx/p/9226078.html
      ii. https://blog.csdn.net/linkequa/article/details/101271096
      iii. https://blog.csdn.net/xzy565143480/article/details/111084642
   b. Type 1 error
      i. The value p=0.05 is the probability of rejecting the null hypothesis when it is true, a Type I Error;
      ii. 数据落在置信区间之外的时候，nh为真，这个时候就犯了第一类错误（误报，false negative
   c. Type 2 error
      i. A Type II Error takes place when the null hypothesis should be rejected, but it does not;

ii. 分布实际上不是样本的分布，但我们采用了样本分布，导致第二类错误（漏报），false positive

d. Reducing the probability of a Type I Error automatically increases the probability of a Type II Error;

e.

| | Positive (1) | Negative (0) |
|---|---|---|
| Positive (1) | True Positive | False Positive (Type I Error) |
| Negative (0) | False Negative (Type II Error) | True Negative |

f.

| | True state of the world | |
|---|---|---|
| Decision | NH True | NH False |
| Reject NH | $p = \alpha$ | $p = 1 - \beta$ |
| Not Reject NH | $p = 1 - \alpha$ | $p = \beta$ |

# Lecture 4: Chi Square

21 October 2021 22:29

## Goodness of fit testing

1.
- Research Hpothesis: The mice learn that the food is on the right and tend to select alleys that go in such a direction;
- Null Hypothesis: The mice select randomly one of the alleys.

2.

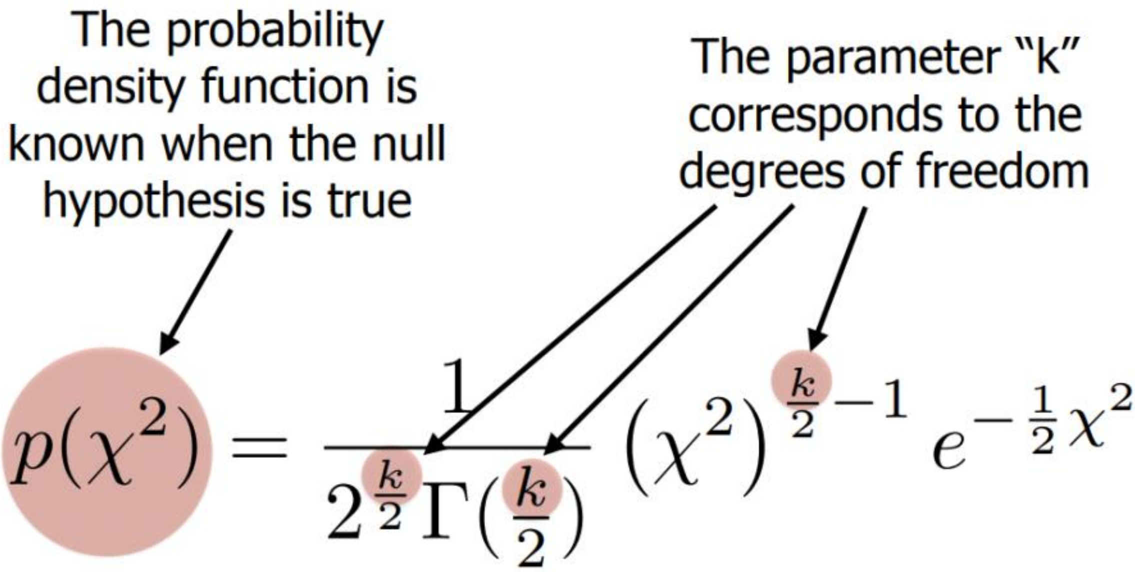
| Alley Chosen | A | B | C | D |
|---|---|---|---|---|
| Observed (O) | 4 | 5 | 8 | 15 |
| Expected (E) | 8 | 8 | 8 | 8 |

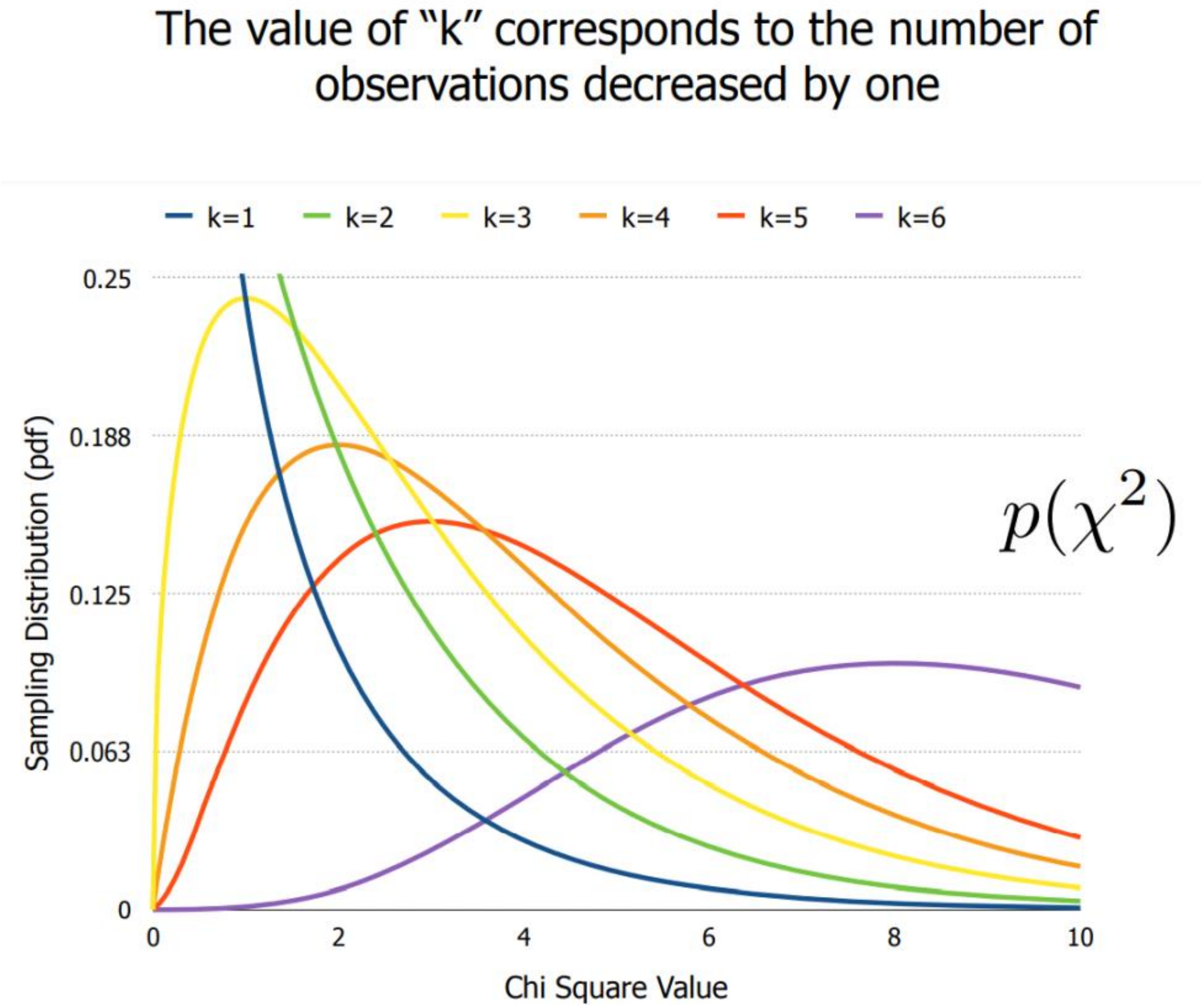
3.

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i}$$

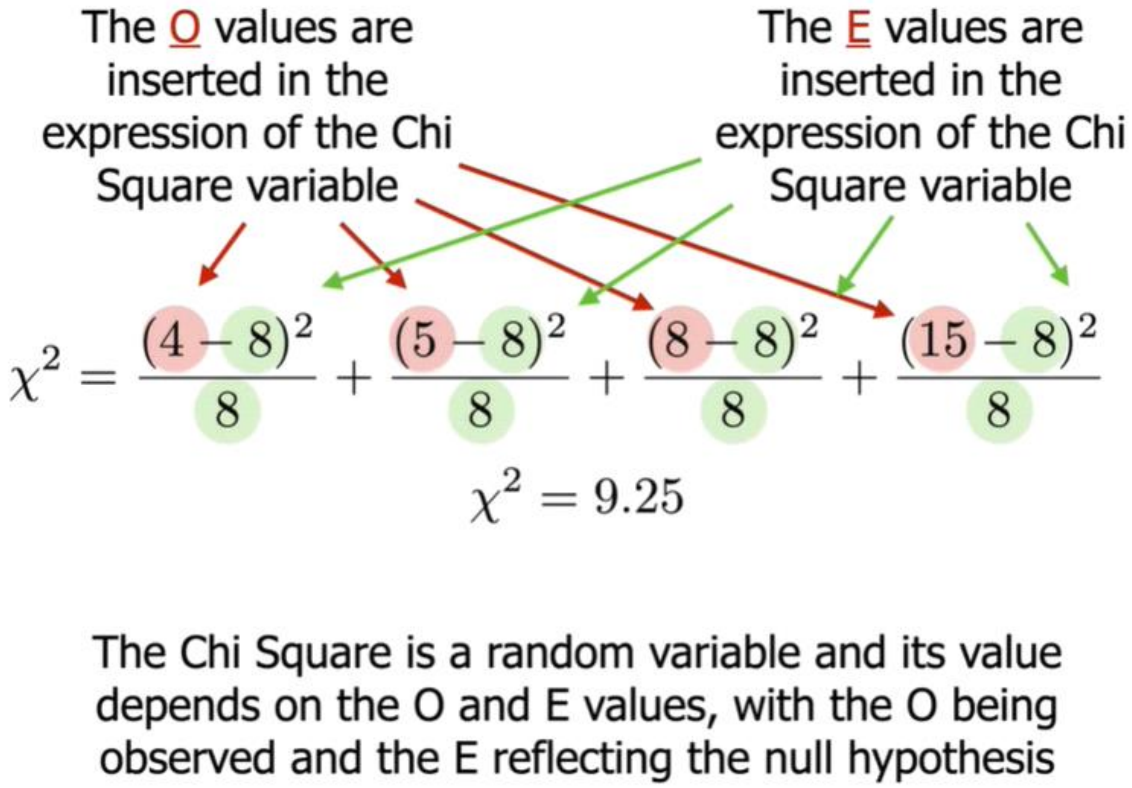
This variable tests whether there is a matching between the observations (O) and the expectations (E)

Sum over all values being compared

4. 


5. 


a. 这里有abcd四个valley，所以N=4， K = N-1 = 3
b. 如果最后得到一个m*n的表格，k= （m-1)*(n-1)

6. 


## Chi square workflow

1. Research hypothesis and null hypothesis
2. Get the observed and expected data
3. Calculate the X^2
4. Calculate the value of k(degree of freedom)
5. Choose the correspond p(x^2) by k，If there are N observations, the number of degrees of freedom is then N-1. 如果最后得到一个m*n的表格，k=（m-1)*(n-1)
6. Plot X^2 on the diagram to see if it can reject the null hypothesis
7.

8.

卡方检验临界值表

| 自由度 | 显著性水平（a） | | | | | |
|---|---|---|---|---|---|---|
| | 0.50 | 0.25 | 0.10 | 0.05 | 0.03 | 0.01 |
| 1 | 0.455 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 |
| 2 | 1.386 | 2.773 | 4.605 | 5.991 | 7.378 | 9.210 |
| 3 | 2.366 | 4.108 | 6.251 | 7.815 | 9.348 | 11.345 |
| 4 | 3.357 | 5.385 | 7.779 | 9.488 | 11.143 | 13.277 |
| 5 | 4.351 | 6.626 | 9.236 | 11.070 | 12.833 | 15.086 |
| 6 | 5.348 | 7.841 | 10.645 | 12.592 | 14.449 | 16.812 |
| 7 | 6.346 | 9.037 | 12.017 | 14.067 | 16.013 | 18.475 |
| 8 | 7.344 | 10.219 | 13.362 | 15.507 | 17.535 | 20.090 |
| 9 | 8.343 | 11.389 | 14.684 | 16.919 | 19.023 | 21.666 |
| 10 | 9.342 | 12.549 | 15.987 | 18.307 | 20.483 | 23.209 |
| 11 | 10.341 | 13.701 | 17.275 | 19.675 | 21.920 | 24.725 |
| 12 | 11.340 | 14.845 | 18.549 | 21.026 | 23.337 | 26.217 |
| 13 | 12.340 | 15.984 | 19.812 | 22.362 | 24.736 | 27.688 |
| 14 | 13.339 | 17.117 | 21.064 | 23.685 | 26.119 | 29.141 |
| 15 | 14.339 | 18.245 | 22.307 | 24.996 | 27.488 | 30.578 |
| 16 | 15.338 | 19.369 | 23.542 | 26.296 | 28.845 | 32.000 |
| 17 | 16.338 | 20.489 | 24.769 | 27.587 | 30.191 | 33.409 |
| 18 | 17.338 | 21.605 | 25.989 | 28.869 | 31.526 | 34.805 |
| 19 | 18.338 | 22.718 | 27.204 | 30.144 | 32.852 | 36.191 |
| 20 | 19.337 | 23.828 | 28.412 | 31.410 | 34.170 | 37.566 |
| 21 | 20.337 | 24.935 | 29.615 | 32.671 | 35.479 | 38.932 |
| 22 | 21.337 | 26.039 | 30.813 | 33.924 | 36.781 | 40.289 |
| 23 | 22.337 | 27.141 | 32.007 | 35.172 | 38.076 | 41.638 |
| 24 | 23.337 | 28.241 | 33.196 | 36.415 | 39.364 | 42.980 |
| 25 | 24.337 | 29.339 | 34.382 | 37.652 | 40.646 | 44.314 |
| 26 | 25.336 | 30.435 | 35.563 | 38.885 | 41.923 | 45.642 |
| 27 | 26.336 | 31.528 | 36.741 | 40.113 | 43.195 | 46.963 |
| 28 | 27.336 | 32.620 | 37.916 | 41.337 | 44.461 | 48.278 |
| 29 | 28.336 | 33.711 | 39.087 | 42.557 | 45.722 | 49.588 |
| 30 | 29.336 | 34.800 | 40.256 | 43.773 | 46.979 | 50.892 |
| 31 | 30.336 | 35.887 | 41.422 | 44.985 | 48.232 | 52.191 |
| 32 | 31.336 | 36.973 | 42.585 | 46.194 | 49.480 | 53.486 |
| 33 | 32.336 | 38.058 | 43.745 | 47.400 | 50.725 | 54.776 |
| 34 | 33.336 | 39.141 | 44.903 | 48.602 | 51.966 | 56.061 |
| 35 | 34.336 | 40.223 | 46.059 | 49.802 | 53.203 | 57.342 |
| 36 | 35.336 | 41.304 | 47.212 | 50.998 | 54.437 | 58.619 |
| 37 | 36.336 | 42.383 | 48.363 | 52.192 | 55.668 | 59.893 |
| 38 | 37.335 | 43.462 | 49.513 | 53.384 | 56.896 | 61.162 |
| 39 | 38.335 | 44.539 | 50.660 | 54.572 | 58.120 | 62.428 |
| 40 | 39.335 | 45.616 | 51.805 | 55.758 | 59.342 | 63.691 |
| 41 | 40.335 | 46.692 | 52.949 | 56.942 | 60.561 | 64.950 |
| 42 | 41.335 | 47.766 | 54.090 | 58.124 | 61.777 | 66.206 |
| 43 | 42.335 | 48.840 | 55.230 | 59.304 | 62.990 | 67.459 |
| 44 | 43.335 | 49.913 | 56.369 | 60.481 | 64.201 | 68.710 |
| 45 | 44.335 | 50.985 | 57.505 | 61.656 | 65.410 | 69.957 |

a. 显著水平通常选为0.05， 1自由度对应值为3.84

b. 比表格中的值大就可以拒绝nh

## Statistical methods for psychology chapter 6

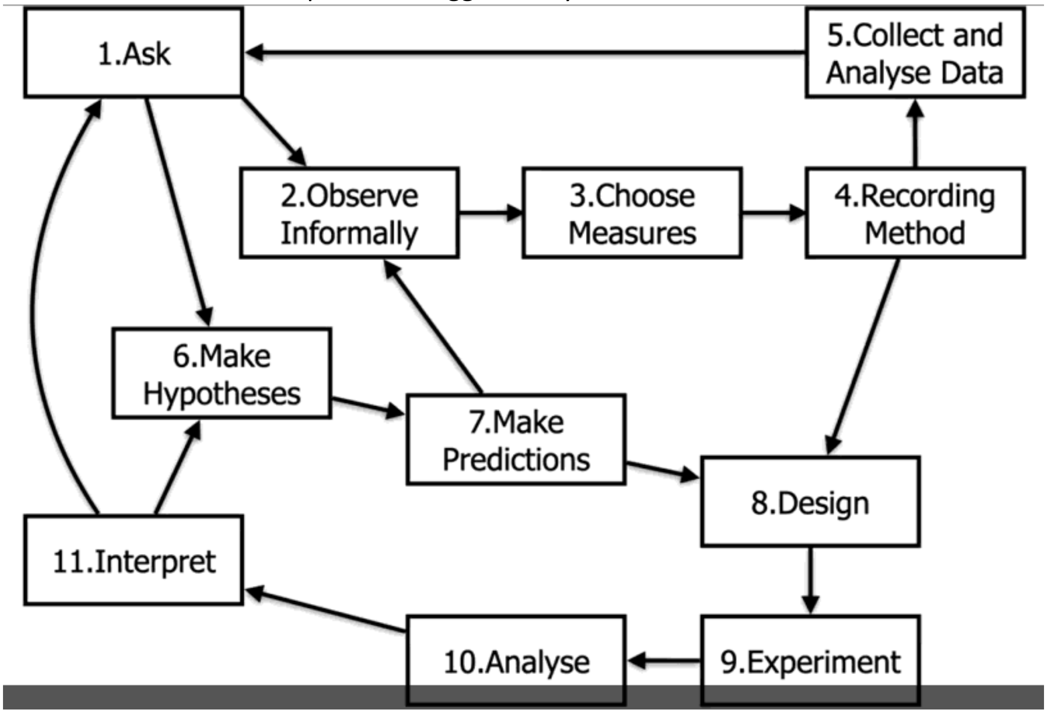
1. Contingency table
   a. https://zhuanlan.zhihu.com/p/41984865

b. https://baike.baidu.com/item/%E5%88%97%E8%81%94%E8%A1%A8/6547006?fr=aladdin

# Lecture 5: observing behavior

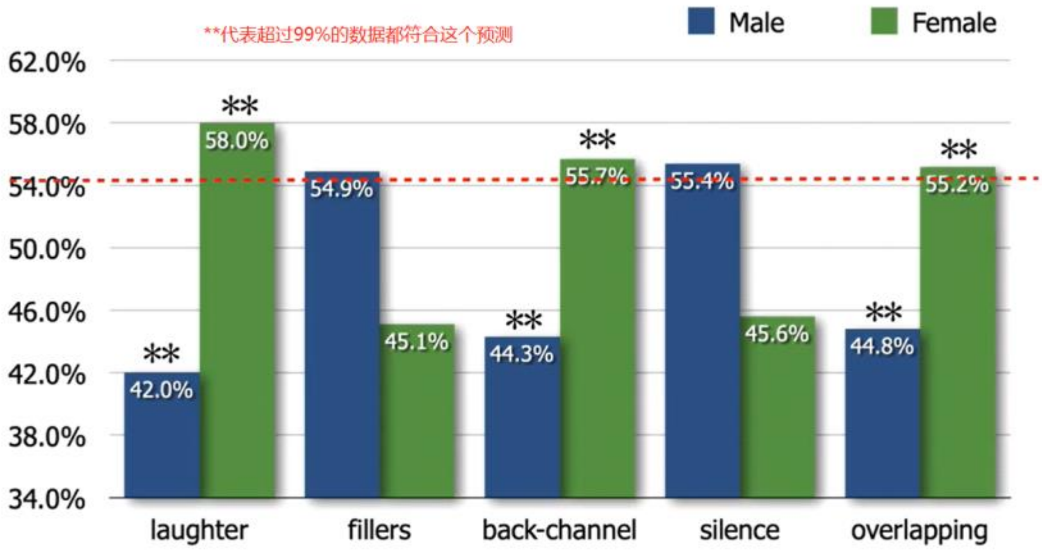
22 October 2021 15:57

## 11 steps of Behavior Analysis

1. Ask a question
2. Make preliminary observation
   a. Deciding what measurements to make
3. Identify the behavioral variables that need to be measured
   a. Choose the variables best for the measurement
4. Choose suitable recording methods
5. Collect and analyze the data
   a. To obtain a clear conclusion with appropriate statistical tools
6. Formulate precise hypothesis
7. Make prediction from the hypothesis
   a. Make more specific predictions that can be tested empirically
8. Design the test
   a. Variables provide the best test of the different predictions
9. Run tests of your hypothesis
10. Analyze the results of your tests
11. Consider alternative interpretations of the evidence
    a. Formulate a list of questions suggested by the data for future research
12. 


## Example

1.



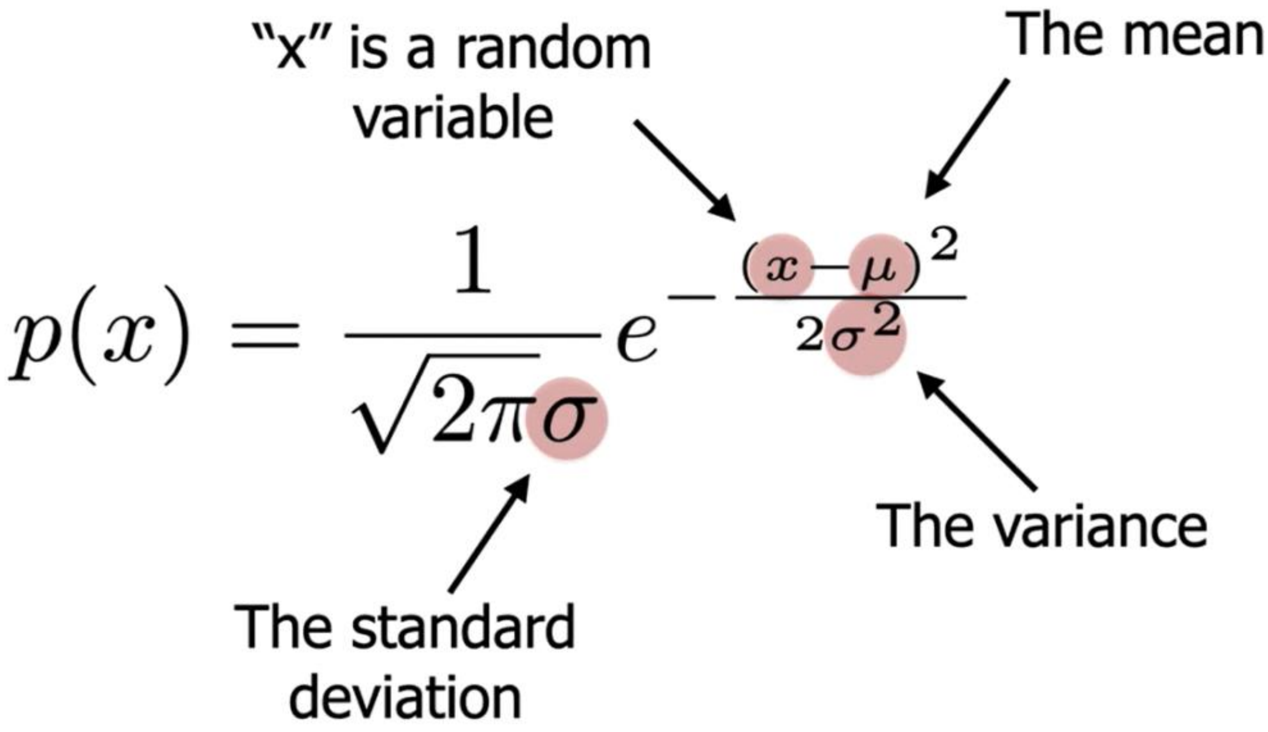
**代表超过99%的数据都符合这个预测
Male
Female
62.0%
58.0%
54.0%
50.0%
46.0%
42.0%
38.0%
34.0%
**
58.0%
54.9%
**
55.7%
55.4%
**
55.2%
45.1%
**
44.3%
45.6%
**
44.8%
**
42.0%
laughter
fillers
back-channel
silence
overlapping

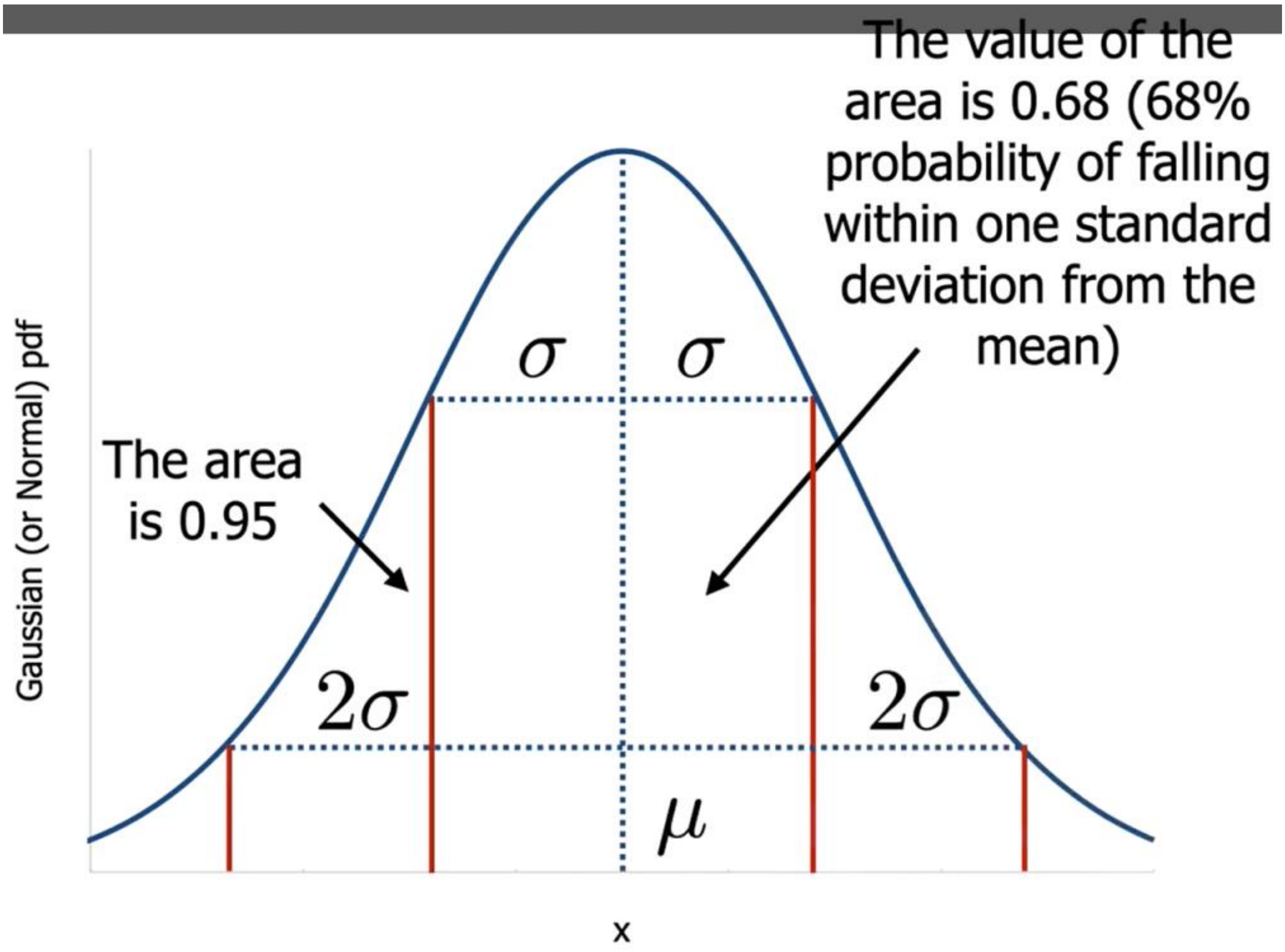

2.

# Lecture 6: students' t

27 October 2021 22:25

## Gaussian distribution



The variable value to be tested (temperature in the toy example)

The values of variable, mean and standard deviation are inserted

$$z = \frac{x - \mu}{\sigma} = \frac{38.0 - 37.0}{0.1} = 10.0$$

The value of z is above the critical value, the null hypothesis can be rejected

The equation can help to find the variable value (temperature) beyond which the null hypothesis can be rejected

$$z \geq 1.65 \Rightarrow \frac{x - \mu}{\sigma} \geq 1.65$$

$$x \geq \mu + 1.65 \cdot \sigma \Rightarrow x \geq 37.165$$

The (toy) temperature threshold beyond which the null hypothesis can be rejected

1. Z变换把变量变换到高斯分布中，获得高斯分布的区间，再解出变量的范围

- It is possible to apply the z-transform and test whether the <u>z value is beyond the critical threshold</u> (1.65 for confidence level 0.05);

## Student's t(t分布)

1. https://zhuanlan.zhihu.com/p/47916928
2. https://www.cnblogs.com/think-and-do/p/6509239.html
3. https://zhuanlan.zhihu.com/p/94429164
4. 例子 https://blog.csdn.net/THUChina/article/details/104980409
5. The Student's t random variable

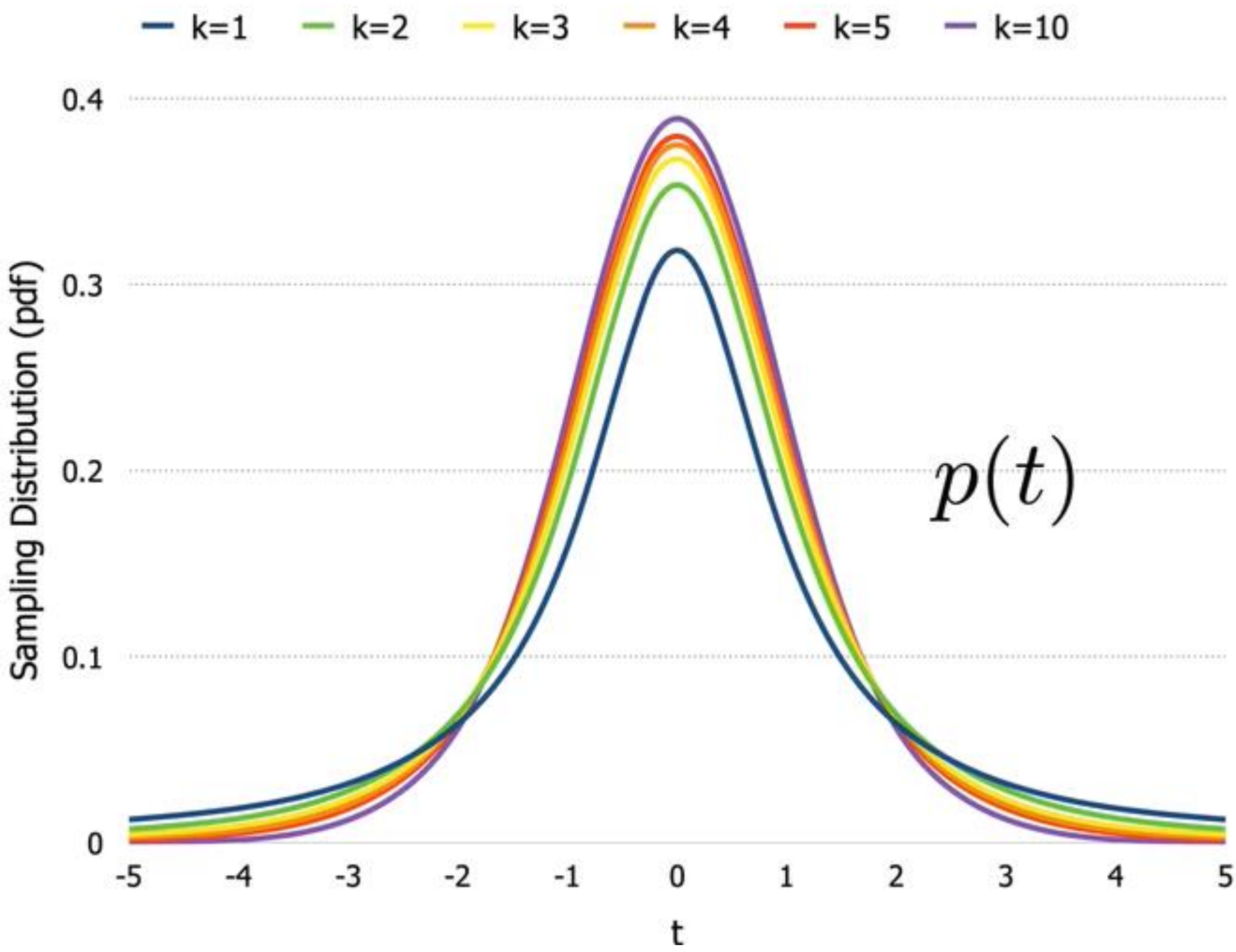
   The sampling distribution of t when the null hypothesis is true is known

   $$t = \frac{\bar{h} - \mu}{\sqrt{\frac{s^2}{N}}}$$

The pdf of t (the sampling distribution) when the Null Hypothesis is true

6. $$p(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{t^2}{k}\right)^{-\frac{k+1}{2}}$$

The parameter k is the number of degrees of freedom

7. 


a. t分布曲线与自由度k有关，k越小，t分布越平坦，中间越低两边越高。k=无限

时，t分布为正态分布

- The Student's t and its sampling distribution allow one to perform hypothesis testing when the mean of the population is known;

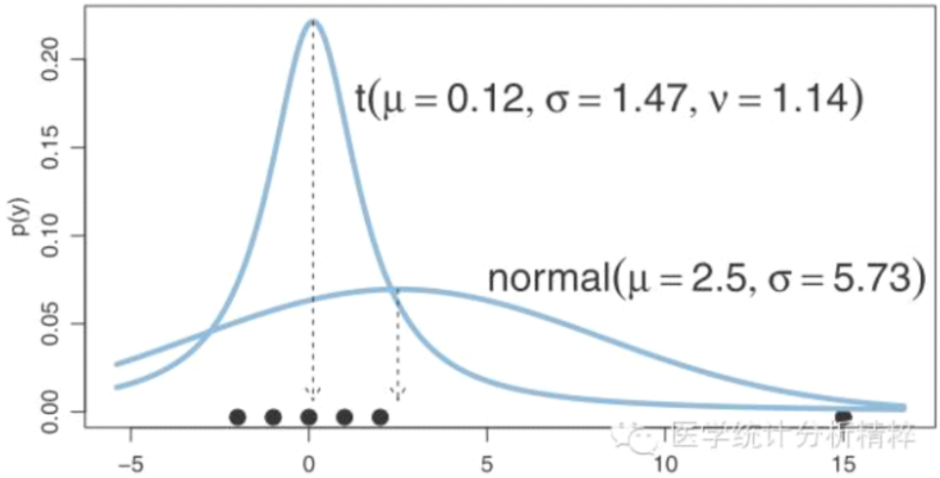
8.

- When t (a function of the data) is beyond the critical value corresponding to a confidence level, it is possible to reject the NH;

9. 正常正态分布容易受到极端值的影响，而t分布可以很好的解决这个问题

我们来比较一下下图中的两条曲线。这两条曲线同样都是对图中底部6个黑色点（数值）进行分布拟合。

我们首先看一下那条矮的、正态分布的曲线。我们前面说过，正态分布的曲线不具备“宽厚”的特征。它的尾部很低，尾部与横轴之间高度很“狭窄”。也就是说，正态分布不能够容忍它长长的尾部出现大概率的事件（图中横轴值为15处一圆点出现概率为六分之一），所以正态分布就很无奈地，将这一点纳入它的胸膛而非留在尾部。于是乎，**恶果**就出现了：图中**正态分布的均数，远远偏离了大多数点所在的位置，标准差也极大**。总之，与我们所期待的很不一致。

a. 


再看一下那条高高的t分布曲线。我们前面说过了，t分布“**温良宽厚**”，它的尾巴很高（本图中不明显，参见上面自由度为1,2,3时所对应的图片），高高的长尾让它有“**容人的雅量**”。所以，这条t分布的曲线，很好的捕捉到了数据点的集中趋势（横坐标：0附近）和离散趋势（标准差：只是那条正态分布曲线标准差的四分之一）。

这也是T分布盛行的原因，即T分布被广泛应用于**小样本**假设检验的原因。虽然是很小的样本，但是，却强大到可以**轻松的排除异常值的干扰，准确把握住数据的特征**（集中趋势和离散趋势）！

10. t分布不需要知道方差，只需要知道均值

## One sample test

1. https://www.jianshu.com/p/a78f4bf79b67
2. https://blog.csdn.net/m0_37777649/article/details/74937242

The random variable is the difference between the means

The mean of the sampling distribution (Central Limit Theorem)

$$\bar{h} = \mu_1 - \mu_2 = \bar{h}_1 - \bar{h}_2$$

3.

$$\sigma^2_{\bar{h}} = \sigma^2_{\bar{h}_1} + \sigma^2_{\bar{h}_2} = \frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}$$

The variance of the difference between the means

The Variance Sum Law

---

Student's t

The Null Hypothesis is that such a difference is null

约等于0

4.

$$t = \frac{\bar{h}_1 - \bar{h}_2 - \boxed{\mu_1 - \mu_2}}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

Student's t

$$t = \frac{\bar{h}_1 - \bar{h}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

5. $$k = (N_1 - 1) + (N_2 - 1)$$

The degrees of freedom is the sum of the individual degrees of freedom

6. Q-Q plot
   a. http://www.360doc.com/content/19/1224/13/68068867_881784673.shtml
   b. https://blog.csdn.net/kdazhe/article/details/104219298
   c. q-q plot是一个散点图。图的横坐标是一组数据的升序分位数（quantile），图的纵坐标是另一组数据的升序分位数。如果这两组数据都是从同一个理论分布中抽样所得，那么这个散点图就会近似于一条直线。如果散点图偏离直线较多，那么我们就可以认为这两组数据不是从同一个理论分布中抽样所得。如果我们只有一组数据并且想检测这组数据是否服从某个理论分布，那么我们的横坐标就选作这个理论分布的分为数的指。当然，我们横坐标的个数应该与纵坐标的个数相同。比如如果我们的样本数据有n nn个，我们想检测这n nn个数据是否是抽样自正态分布，那么我们的横坐标就选取标准正太分布的n nn个分位点。
7.
8.

# Psychometrics

02 November 2021 22:47

## Alphabet

| | |
|---|---|
| Psychometrics | |

# Lecture 7: Psychometrics

02 November 2021 22:47

本章只需要学PPT，文章中的问卷方法需要了解

## Big Five

1. Extraversion
2. Agreeableness
3. Conscientiousness
4. Neuroticism
5. Openness
6. 

| ID | Item | SD | D | NA | A | SA |
|---|---|---|---|---|---|---|
| 1 | I am reserved | | | | | |
| 2 | I am generally trusting | | | | | |
| 3 | I am lazy | | | | | |
| 4 | I am relaxed, I handle stress well | | | | | |
| 5 | I have few artistic interests | | | | | |
| 6 | I am outgoing, sociable | | | | | |
| 7 | I tend to find faults with others | | | | | |
| 8 | I do a thorough job | | | | | |
| 9 | I get nervous easily | | | | | |
| 10 | I have an active imagination | | | | | |

7. 
   - Openness: Item 10 - Item 5
   - Conscientiousness: Item 8 - Item 3
   - Extraversion: Item 6 - Item 1
   - Agreeableness: Item 2 - Item 7
   - Neuroticism: Item 9 - Item 4

## 6-Steps of Scale Development

1. Item Generation
   a. Deductive
   b. Inductive
2. Questionnaire Administration
   a. Test the sample
3. Initial Item Reduction
4. Confirmatory Factor Analysis
5. Convergent/Discriminant Validity
6. Replication

## Vocabulary

| | |
|---|---|
| Latent | 潜在的 |
| phenotypic individual differences | 表型独立差异 |
| Assertive | 自信的 |
| Neuroticism | 神经过敏 |
| Conscientiousness | 责任心 |
| Agreeableness | 宜人性 |
| Touchy | 敏感的，难以取悦的 |
| Reserved | 寡言少语的 |
| parsimonious | 吝啬的，确切的 |
| Convergent | 收敛 |
| attenuated | 衰减 |

# Lecture 8: Relationships

02 November 2021 22:47

## Covariance

1. https://www.cnblogs.com/leezx/p/9929340.html

The covariance of X and Y

The average value of X

The average value of Y

2. $$\sigma_{xy} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y - \bar{y})}{N - 1}$$

The total number of pairs (x,y)

## Regression

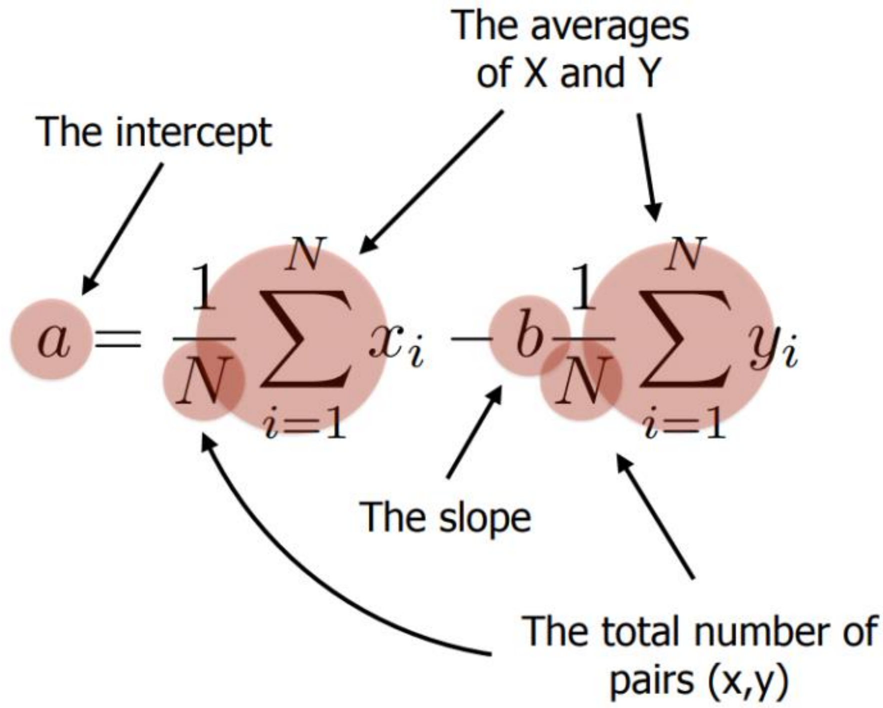
1. https://www.jianshu.com/p/02cc8c8f49af?utm_source=desktop

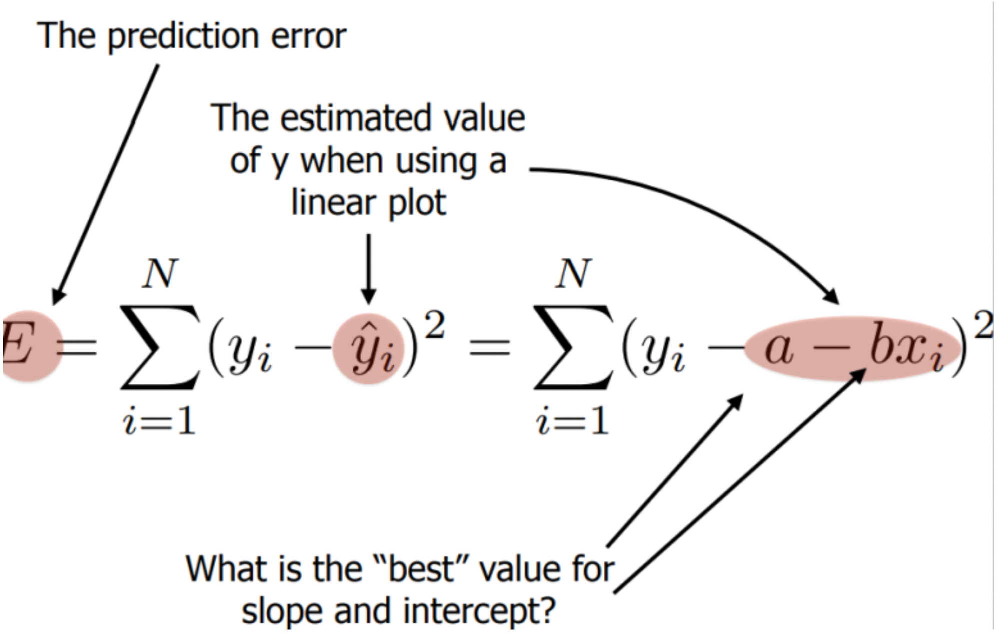
The slope

The covariance of X and Y

2. $$b = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} = \frac{\sigma_{xy}}{\sigma_x^2}$$

The variance of X

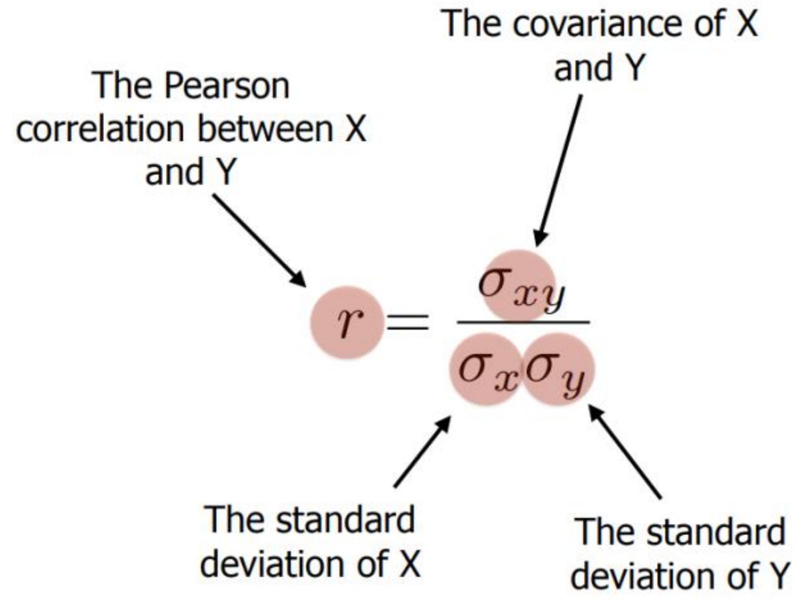
3. 


4. 


5. 其实就是线性回归，E就是损失函数

## Correlation

1. https://blog.csdn.net/ruthywei/article/details/82527400
2. https://blog.csdn.net/chao2016/article/details/80917579

3. 


4. 皮尔森相关系数反映了两个变量的线性相关性的强弱程度，r的绝对值越大说明相关性越强。

   - 当r>0时，表明两个变量正相关，即一个变量值越大则另一个变量值也会越大；
   - 当r<0时，表明两个变量负相关，即一个变量值越大则另一个变量值反而会越小；
   - 当r=0时，表明两个变量不是线性相关的（注意只是非线性相关），但是可能存在其他方式的相关性（比如曲线方式）；
   - 当r=1和-1时，意味着两个变量X和Y可以很好的由直线方程来描述，所有样本点都很好的落在一条直线上。

5. 
   - It is possible to measure the relationship between two variables, i.e., their tendency to change according to each other;
   - The relationship has no direction (it is not possible to say whether one variable influences the other or vice versa);
   - However, the analysis of the relationships can provide insight about the phenomena under exam.

## Vocabulary

| | |
|---|---|
| Covariance | 协方差 |
| experimenter | 实验者 |

| | |
|---|---|
| Predictor | 预测者 |
| Intercept | 截距 |
| Slope | 坡度 |
| Correlation | 相关 |
| Standard deviation | 标准差 |
| | |

# Lecture 9: Judgment's Studies

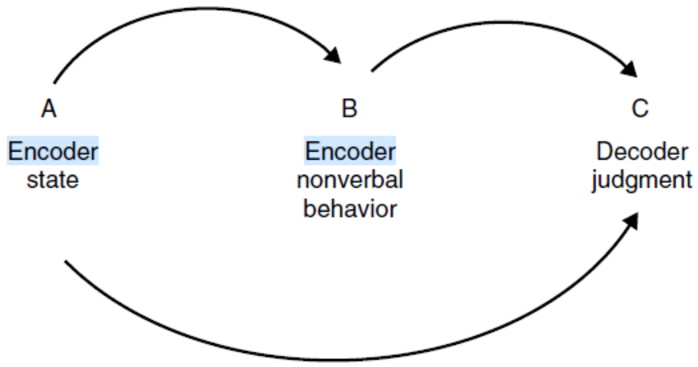
02 November 2021 22:47

## Introduction

1. Judgement studies refers to those studies in which behaviors, persons, objects or concepts are evaluated by one or more judges.
2. Type
   a.

| Dimensions | Examples |
|---|---|
| Type of Variable | Dependent vs Independent |
| Measurement Units | Physical vs Psychological |
| Reliability | Lower vs Higher |
| Social Meaning | Lower vs Higher |

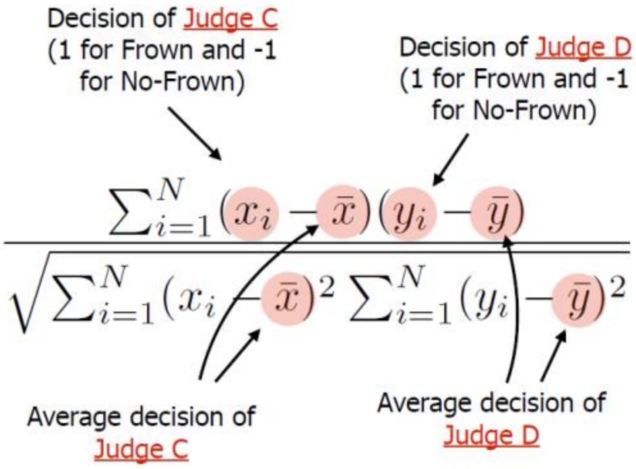
## The judgement study model

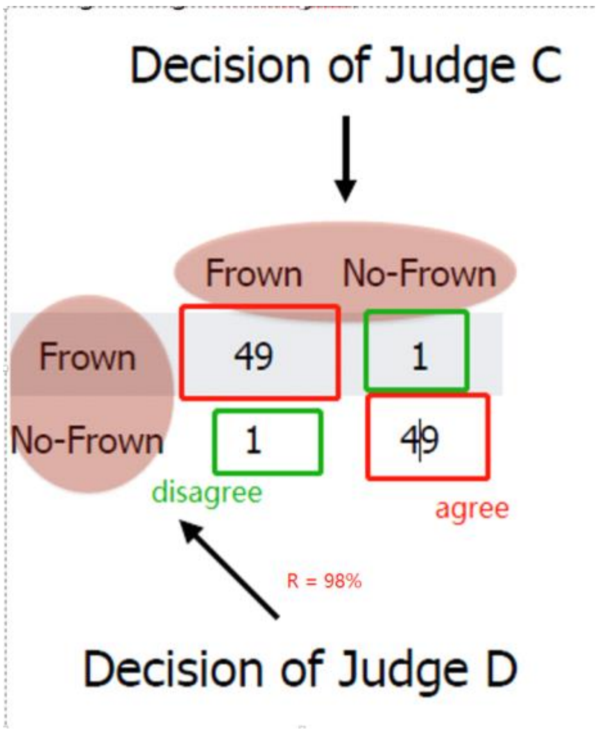



1. Encoder state(A)
   a.
2. Encoder nonverbal behavior(B)
3. Decoder judgement (C)

## Main issue: reliability

1. Reliability: how reliable are the judgements
   a. There is consistency between observations and judgements
   b. The reliability can be thought of as the measure of the consensus among multiple judges
   c. The higher the consensus, the higher the reliability
2. Selection: how to select judges
3. Composition: how to combine judgements to form composite variables
4. Reliability benchmark
   a. Percentage of agreement
      i. R = Agree/(Agree+Disagree)*100%
      ii. r =

1) 


iii. 

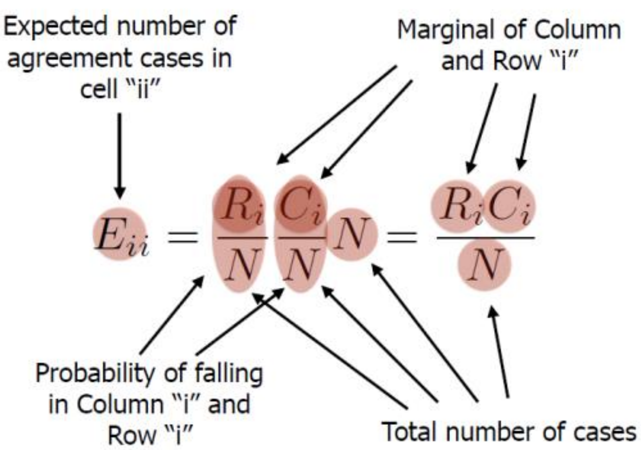

iv. It can be high as there is no variance in the judgements

v. The same value of R correspond to different values of correlation

b. Kappa value

i. K = (Observed - Expected)/(Agree + Disagree - Expected)

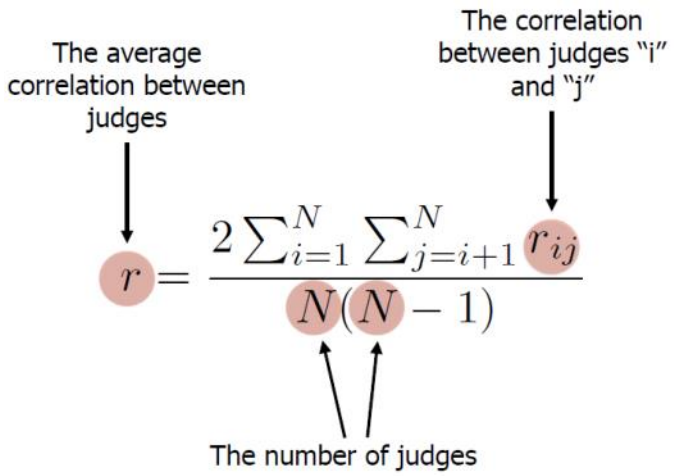
ii. E

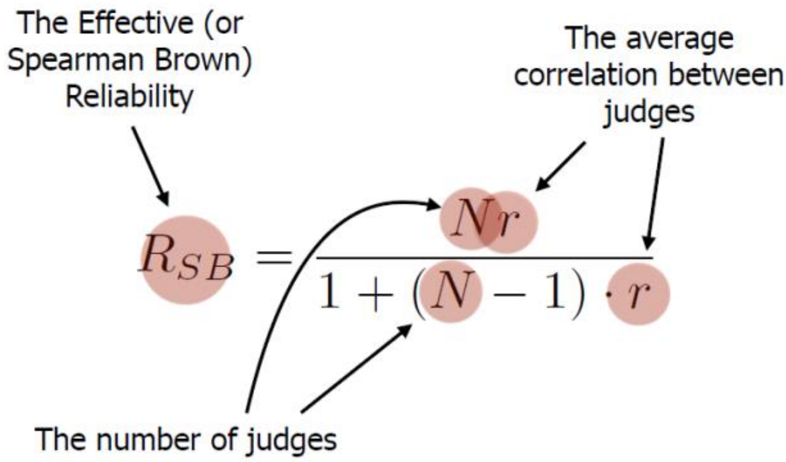
1) 


iii. Judgements are random

iv. Not suitable when more than two categories(2x2)

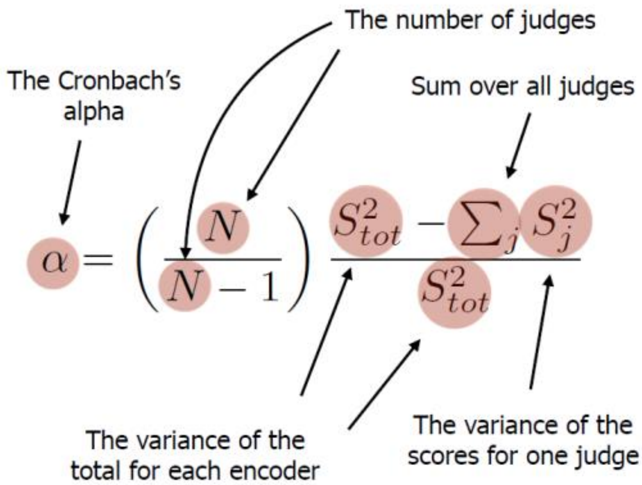
c. Effective(Spearman Brown) reliability

i. 


ii. 


iii. How much associated are the judgments of two random judges
iv. It is an average value
v. It does not say whether all judges are equally correlated with one another

d. Cronbach's alpha

i. 


ii. It avoids the calculation of multiple correlations when the number of judges is high;
iii. It tends to give the same values as the other reliability measures considered in this lecture (it is affected by the same limitations).

## Conclusion

1. Judgement studies allow one to answer questions on how inner states are expressed and perceived

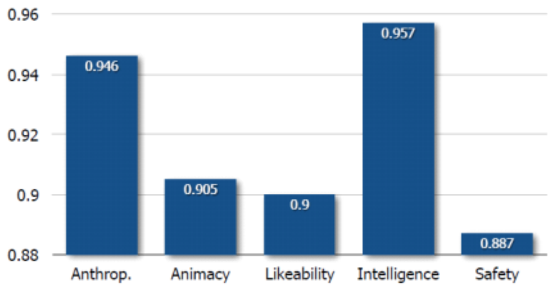
2. Reliability measures are expected to quantify the extent to which multiple judgements agree with one another
3. Multiple judges agree is not sufficient, they must have high mutual correlation

# Vocabulary

# Lecture 10: synthetic impressions I

13 May 2022 12:29

## Synthetic impressions

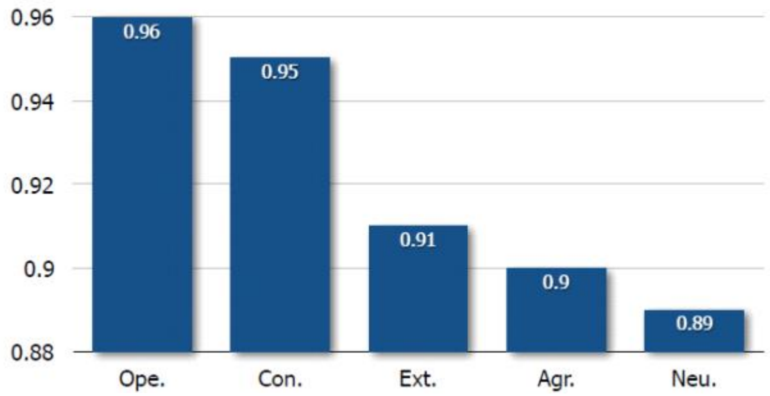
1. The godspeed questionnaire
    a. Compares the results from different studies of five key concepts in HRI
2. Anthropomorphism
    a. Attributions of a human form to nonhuman things
3. Animacy
    a. The classic perception of life
4. Likability
    a. Positive first impressions often lead to more positive evaluations
5. Perceived intelligence
    a. perceived intelligence of a robot will depend on its competence
6. Perceived safety
    a. Describes the user's perception of danger when interacting with a robot
7. Reliability
    a. 


## Gestures and godspeed scores

1. Motivation
    a. Gestures are used to communicate
2. The setting
    a. 30 observers filled godspeed questionnaire and Big-Five rating the gestures
3. Calculation
    a. $$\chi^2 = \sum_{\alpha} \sum_{\lambda} \frac{(c_j^{(\alpha,\lambda)} - E)^2}{E}$$
4. Conclusion
    a. Find the relationship between gesture and the perception of users
    b. Animacy and Likability determins whether there is more interaction
    c. The core gesture “Pointing” does not show any interaction between shape and perception

## Gestures and personality

1. The big five traits
    a. Describe most phenotypic individual differences
    b. Extraversion, agreeableness, conscientiousness, neuroticism, openness
    c. Reliability

i. 


d. $$\chi^2 = \sum_{\alpha} \sum_{\lambda} \frac{(t_j^{(\alpha,\lambda)} - E)^2}{E}$$

2. Conclusion
    a. Find the relationship between gesture and big five traits
    b. There is a significant interplay between godspeed score and big-five
    c. It is possible to change the perception of the users by changing the personality impressions that the robots convey

## Conclusion

1. Gesture give rise to a wide spectrum of synthetic impressions
2. The impressions follow laws in human-human interactions
3. Collect data in real-world setting

# Lecture 11: synthetic impressions II

13 May 2022 12:30

## Synthetic impressions

1. recap

## Gestures and the attraction paradigm

1. The Attraction Paradigm
   a. evaluations of strangers appear to be affected by degree of similarity
   b. perceived similarity predicted attraction
2. Formula
   a. $$r = 1 - \frac{6\sum_{k=1}^{M} d(d_k, g_k)}{M(M^2 - 1)}$$
3. Recap
   a. The attraction paradigm can be exploited effectively only if it is possible to understand which of the effects is taking place.

## Gestures and understandability

1. Understandability
   a. Gestures are linked to a meaning
2. Interpretation
   a. Rate 10 possible interpretations of every gesture
3. Calculating the understandability and the correlations
4. Recap
   a. There is an association between changes in amplitude and changes in understandability;
   b. There is a statistically significant correlation between understandability and Godspeed scores;
   c. The interplay appears to reproduce the incompatibility between social and task skills

## Conclusions

1. Gestures give rise to a wide spectrum of synthetic impressions
2. The impressions appear to follow the rules in human-human interactions

# Behaviour Understanding

12 May 2022 15:18

# Lecture 12: Bayesian decision theory

13 May 2022 12:37

## Introduction

1. The Feature vectors
   a. $\vec{x} = (x_1, \dots, x_j, \dots, x_D)$
   b. D features in total
2. The decisions
   a. $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_i, \dots, \mathcal{C}_M\}$
   b. M decisions in total
   c. In AI, a decision means to map a feature vector into a decision, map X to C
   d.

## Prior rule

1. https://zh.wikipedia.org/wiki/%E5%85%88%E9%A9%97%E6%88%87%E5%BE%8C%E9%A9%97
2. $\mathcal{C}^* = \arg\max_{\mathcal{C}_k \in \mathcal{C}} p(\mathcal{C}_k)$

## Posterior rule

1. https://zh.wikipedia.org/wiki/%E5%85%88%E9%A9%97%E6%88%87%E5%BE%8C%E9%A9%97
2. $p(\mathcal{C}_i|\vec{x}) = \dfrac{p(\vec{x}|\mathcal{C}_i)p(\mathcal{C}_i)}{p(\vec{x})}$

## Bayesian decision theory

1. Example, P474
2. 得到先验priors
3. 得到likelihoods
4. 求出每个action对应的posterior, 选取最大的那个

## Example

# Lecture 13: Discriminant function

13 May 2022 12:31

## Conditional risk and bayes decision rule

1. Loss
   a. The loss resulting from performing action "i" when the decision is "j"

      A possible action

      A possible decision

      Number of possible actions

      $$\pi(\mathcal{B}_i|\mathcal{C}_j)$$
      $$i = 1, \ldots, T$$
      $$j = 1, \ldots, M$$

      Number of possible decisions
2. Conditional risk
   a. Expected loss (or conditional risk) associated to the action

      A possible action

      $$\mathcal{R}(\mathcal{B}_i|\vec{x}) = \sum_{j=1}^{M} \pi(\mathcal{B}_i|\mathcal{C}_j)p(\mathcal{C}_j|\vec{x})$$

      Sum over all possible classes the vector can be assigned to

      The posterior of the class acts as a weight in the sum
   b. The conditional risk is the weighted sum of the losses associated to an action
   c.
3. Bayes decision rule
   a. The Bayes Decision Rule targets the decision that minimises the conditional risk.

## Binary classification

1. Two classes and two actions
   a. For example, the classes are "healthy" and "ill" while the actions are "keep in the hospital" and "do not keep in the hospital".
2. Expected loss/ conditional risk
   a. $$\mathcal{R}(\mathcal{B}_1|\vec{x}) = \pi(\mathcal{B}_1|\mathcal{C}_1)p(\mathcal{C}_1|\vec{x}) + \pi(\mathcal{B}_1|\mathcal{C}_2)p(\mathcal{C}_2|\vec{x})$$
      $$\mathcal{R}(\mathcal{B}_2|\vec{x}) = \pi(\mathcal{B}_2|\mathcal{C}_1)p(\mathcal{C}_1|\vec{x}) + \pi(\mathcal{B}_2|\mathcal{C}_2)p(\mathcal{C}_2|\vec{x})$$
   b. Minimize the loss by difference

$$\mathcal{R}(\mathcal{B}_2|\vec{x}) - \mathcal{R}(\mathcal{B}_1|\vec{x}) =$$

i. $$= [\pi(\mathcal{B}_2|\mathcal{C}_1) - \pi(\mathcal{B}_1|\mathcal{C}_1)]\, p(\mathcal{C}_1|\vec{x}) + \\ + [\pi(\mathcal{B}_2|\mathcal{C}_2) - \pi(\mathcal{B}_1|\mathcal{C}_2)]\, p(\mathcal{C}_2|\vec{x})$$

ii. Get the minimize when the difference is positive

c. Minimize the loss by ratio

i. $$\frac{p(\vec{x}|\mathcal{C}_1)}{p(\vec{x}|\mathcal{C}_2)} > \frac{\pi(\mathcal{B}_1|\mathcal{C}_2) - \pi(\mathcal{B}_2|\mathcal{C}_2)}{\pi(\mathcal{B}_2|\mathcal{C}_1) - \pi(\mathcal{B}_1|\mathcal{C}_1)} \frac{p(\mathcal{C}_2)}{p(\mathcal{C}_1)}$$

ii. Get the minimize when the likelihood ratio satisfies the equation above

3. Example

## Zero-one loss and posterior rule

1. Zero one loss

a. $$\pi(\mathcal{B}_i|\mathcal{C}_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases}$$

The expected loss (conditional risk) when performing action "i"

The sum includes only the cases in which "j" and "i" are different

a. $$\mathcal{R}(\mathcal{B}_i|\vec{x}) = \sum_{j \neq i} \pi(\mathcal{B}_i|\mathcal{C}_j) p(\mathcal{C}_j|\vec{x}) = \\ = 1 - p(\vec{C}_i|\vec{x})$$

The highest posterior minimises the conditional risk

2. Generally, The number of actions is the same as the number of decisions
3. Generally, for every class one decision is right while all the others are wrong
4. Generally, the minimisation of the conditional risk is equivalent to the application of the posterior rule

## Gaussian discriminant function

1. $$\mathcal{G} = \{\gamma_1(\vec{x}), \ldots, \gamma_M(\vec{x})\}$$

2. $$\gamma_k(\vec{x}) = -\mathcal{R}(\mathcal{B}_k|\vec{x}) = p(\mathcal{C}_k|\vec{x}) - 1 \\ \gamma_k(\vec{x}) \simeq p(\mathcal{C}_k|\vec{x})$$

a. γ为损失函数, 约等于后验
b. γ是risk的负数, 因此γ越大, risk越小

3. $$\log \gamma_k(\vec{x}) = \log \frac{p(\vec{x}|\mathcal{C}_k)p(\mathcal{C}_k)}{p(\vec{x})}$$

$$\log \gamma_k(\vec{x}) \simeq \log p(\vec{x}|\mathcal{C}_k) + \log p(\mathcal{C}_k)$$

a. 取对数为了把乘法变成加法, 方便计算

4. $$\hat{k} = \arg \max_{k\in[1,M]} \log p(\vec{x}|\mathcal{C}_k) + \log p(\mathcal{C}_k)$$

a. k为decision, 因为是负数, 所以取最大值

5. $$\log p(\vec{x}|\mathcal{C}_k) = \sum_{i=1}^{D} \log p(x_i|\mathcal{C}_k)$$

6. $$p(x_i|\mathcal{C}_k) = \frac{1}{\sqrt{2\pi}\sigma_{ik}} \exp\left[-\frac{(x_i - \mu_{ik})^2}{2\sigma_{ik}^2}\right]$$

Sum over the standard deviations of the individual features

True when features statistically independent given the class

7. $$\log p(\vec{x}|\mathcal{C}_k) = -\sum_{i=1}^{D}\left[\log \sqrt{2\pi}\sigma_{ik} + \frac{(x_i - \mu_{ik})^2}{2\sigma_{ik}^2}\right]$$

Euclidean distance between feature vector and class average

# Lecture 14: Training and test

13 May 2022 14:32

## Training of the likelihood

## Training of the priors

## k-fold and performance measurement

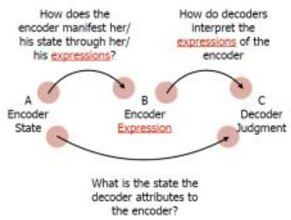
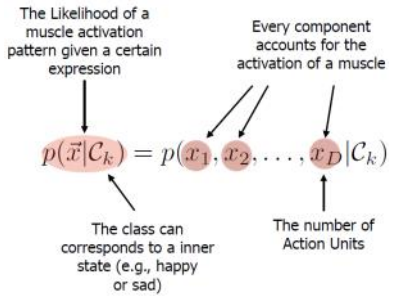
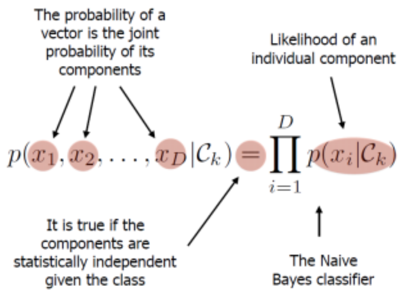
## Conclusion

# Lecture 15: Facial expression

13 May 2022 12:30

## Nonverbal communication

## Facial expression

1. Model
   a. 
   
2. Facial action code(FAC)
   a. 
   
   b. 
   

## Action units (AU)

1. There are 46 action units on the face
2. An AU corresponds to the activation of one or more muscles
3. A facial expression is a combination of Aus
4. Different facial expressions convey different social messages

## Facial expression analysis

1. 

## Conclusion

1. There is a gap between the AU being displayed and the interpretation of an expression
2. Machines can automatically detect Aus

# Lecture 16: Basic signal processing I

13 May 2022 14:35

## Introduction

1. A signal is a measurable quantity changes over time and it is continuous
2. A digital signal is a sequence of signals collected at fixed time step, it is discrete
3. We will never know what happens in the signal between two consecutive samples

## Time domain processing

1. Formula

   The convolution between the signal "s" and the window "w"

   The sum extends over all samples of the signal

   a. $$y[k] = \sum_{n=-\infty}^{\infty} s[n]w[k-n]$$

   The window

   The signal

2. The rectangular window

   a. $$w[n] = \begin{cases} 1 & : \quad 0 \leq n \leq N-1 \\ 0 & : \quad n < 0 \\ 0 & : \quad n > N-1 \end{cases}$$

3. Convolution between the signal s and window w

   a. $$y[k] = \sum_{n=k-N+1}^{k} s[n]w[k-n]$$

4. The hamming window

   a. $$w[n] = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) & : \quad 0 \leq n \leq N-1 \\ 0 & : \quad n < 0 \\ 0 & : \quad n > N-1 \end{cases}$$

## Short-term analysis

1. The short-term properties of a signal can be calculated through the convolution with an analysis window;
2. The Energy

   $$f(s[n]) = \frac{(s[n])^2}{N}$$

   $$f(s[n]) = \frac{|s[n]|}{N}$$

   The Magnitude

## Conclusions

1. Speech can be analysed in the time domain through convolution operations
2. Process take place in the freq domain
3. Speech is the main form of communication

# Lecture 17: Basic signal processing II

13 May 2022 14:34

## Zero crossing rate (ZCR)

1. Formula

   The Zero Crossing Rate

   a. $$Z[k] = \sum_{n=-\infty}^{\infty} \frac{g(s[n], s[n-1])}{2N} w[k-n]$$

   The ZCR is half the number of times the signal crosses the horizontal axis in the window

2. The ZCR when the signal is a sinusoid of frequency "f"

   The sinusoid frequency

   $$Z[k] = \frac{2f}{F}$$

   The sampling frequency

## Autocorrelation

1. Formula

   The parameter "m" is called the lag

   The Autocorrelation

   a. $$R_m[k] = \sum_{n=-\infty}^{\infty} s[n]w[k-n]s[n+m]w[k-n-m]$$

   Product of two samples at distance "m" from one another

2. The ZCR provides information about the frequency that carries most energy
3. The autocorrelation changes depending on how periodic is the signal
4. The inverse of the distance between the maxima of the autocorrelation accounts for the frequency that carries most energy

## Fourier transform

1. Discrete fourier transform

   a. $$X[k] = \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k}{N} n}$$

2. DFT and its inverse IDFT represent the same information in two different ways
3. The spectrogram shows the energy distribution across the frequencies
4. The DFT changes depending on the property of the signal

## Conclusion

# Lecture 18: Speech personality

13 May 2022 14:35

## Introduction

1. The short-term properties of the speech signals provide information about the speaking style (how people speak);
2. The speaking style influences the impression people develop about the speaker;
3. It is possible to infer the personality traits people attribute to a speaker from her/his speaking style.

## Computational paralinguistics

1. Short-term features
   a. Big three of prosody: energy, pitch and tempo
   b. Energy and magnitude account for energy
   c. Zero crossing rate and autocorrelation accounts for pitch
   d. The number of syllables per minutes accounts for tempo

## Trait prediction

1. Training
   a. Aims at finding the parameters(theta and threshold)
   b. The training is performed by minimize the error rate
   c. k-fold approach
2. Speech and personality
   a. Two dimensions underlie most judgements of traits
      i. People, reference to competence
      ii. Cultures, reference to warmth
   b. Rate and pitch variation were the most influential for competence and benevolence respectively

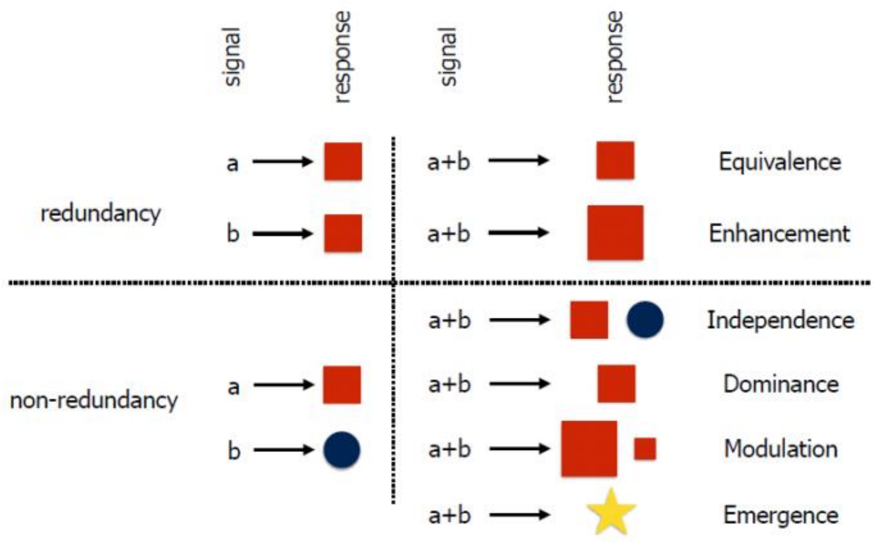
# Lecture 19: Multimodal

12 May 2022 15:18

## Multimodality (Psychology & Neuroscience): Gestalt Theory

1. Gestalt Theory describes different laws or principles for perceptual grouping of information into a coherent whole, including proximity, symmerty, similarity and closure
    a. https://en.wikipedia.org/wiki/Gestalt_psychology#Law_of_symmetry
    b. Proximity: Objects close to each other tend to be considered grouped together, 离得近的被理解为一组
    c. symmetry： More symmetric interpretations tend to be favoured, 倾向于理解为对称的东西
    d. Similarity: similar objects tend to be considered grouped together, 相似的被理解为一组
    e. Closure: close and complete figures tend to be perceived in absence of continuous lines, 残缺的倾向于被理解成完整的
    f. Continuity: lines that follow one another tend to be considered continuous, 两个重叠的对象, 重叠的部分倾向于被理解成连贯的, 而不是分开的两块

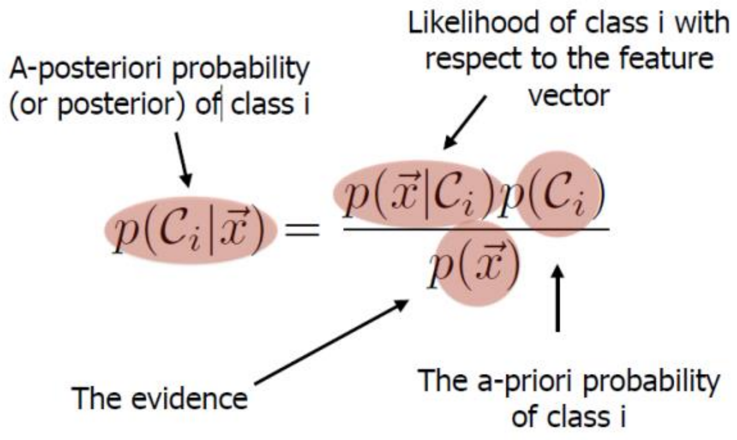
## Multimodality (Communication & Life Science)

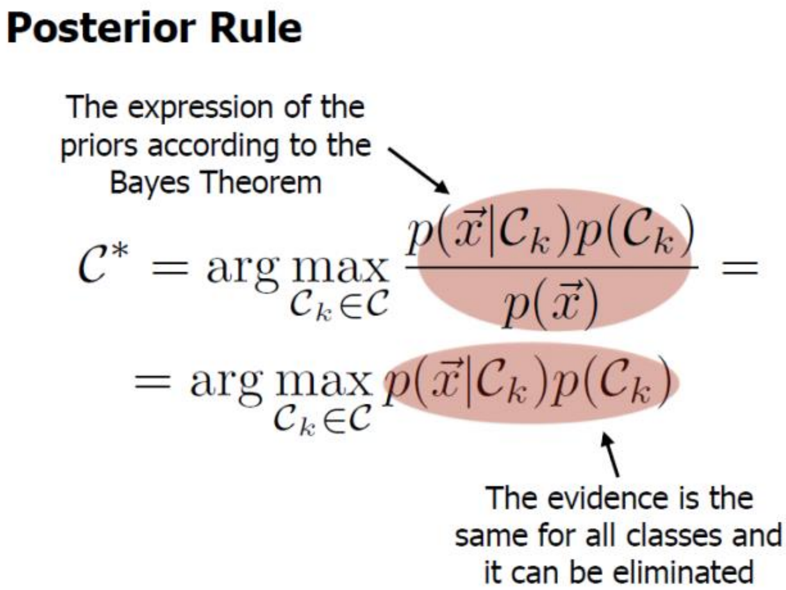
1. Lecture 19-text2
    a. 


Partan and Marler, "Issues in the Classification of Multimodal Communication Signals", The American Naturalist, 166(2):231-245, 2005.

## Multimodality (Computing Science & AI)

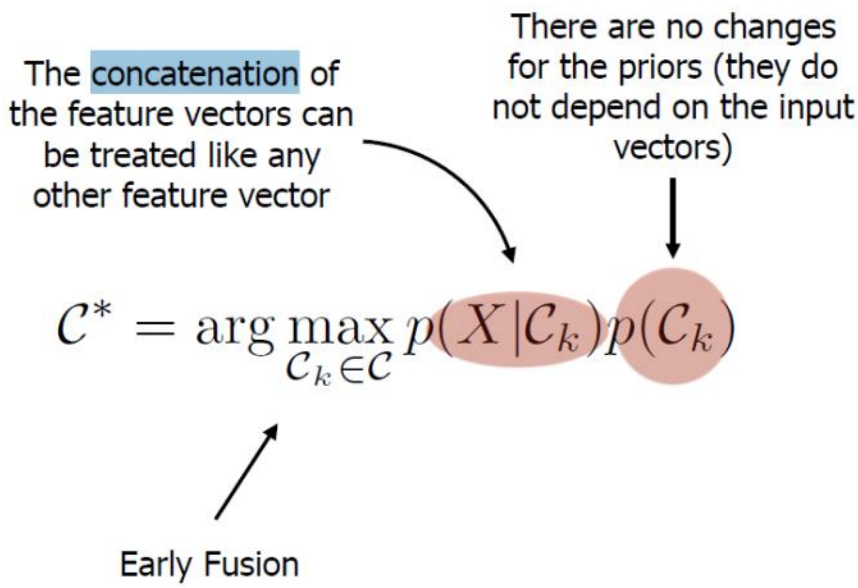
1. The bayes theorem

a. 


b. 

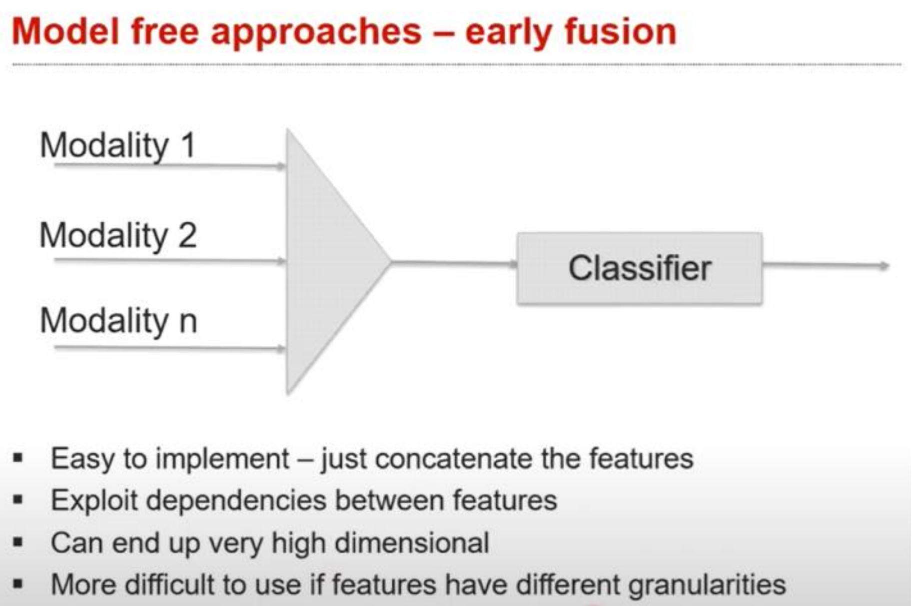

2. Early fusion

Early and late fusion are not expected to explain the way living beings respond to multimodal
communication patterns. They are simply methodologies that allow a machine to map a multimodal input pattern X into a suitable response ^!. In other words, early and late fusion
can maybe reproduce the observations summarised in Figure 1.2, but they cannot be considered
an explanation or a model of the processes that lead from stimulus to response in living beings.

a. 


b.
## Recap

- The early fusion is the concatenation of the feature vectors extracted from the data captured through multiple sensors;
- The concatenation can be treated like any other vector;
- In the early fusion case, there are no changes from a decision theoretic point of view.

c. CMU
   i. https://www.youtube.com/watch?v=2xr4P0WGKSA
   ii. Classifier do the fusion
   iii. 

3. Late fusion

There are no changes for the priors (they do not depend on the input vectors)

a. $$C^* = \arg\max_{C_k \in C} p(\vec{x}_1, \ldots, \vec{x}_R | C_k) p(C_k)$$

The likelihood must be changed to reflect the presence of multiple feature vectors

The assumption is that the input vectors are statistically independent given the class

Product over all sensors

b. $$p(\vec{x}_1, \ldots, \vec{x}_R | C_k) = \prod_{j=1}^{R} p(\vec{x}_j | C_k)$$

If one term is close to zero, the entire product is close to zero

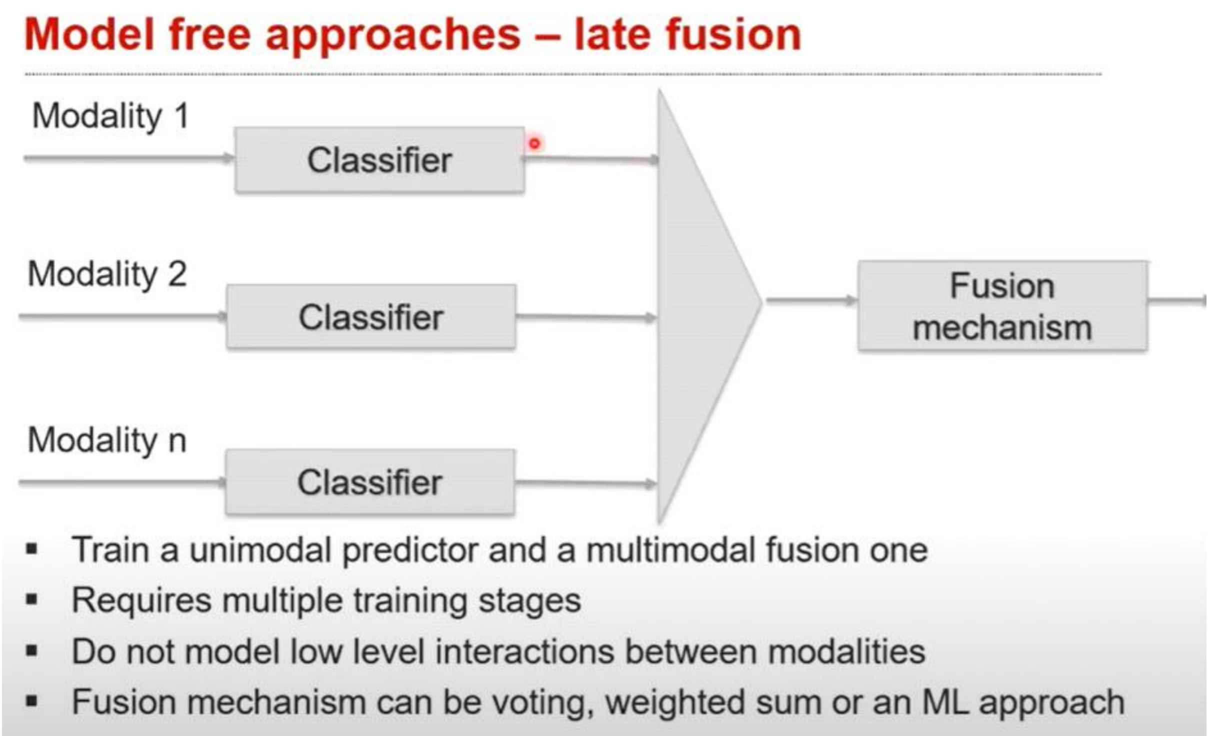
c.

## Recap

- The late fusion is the combination of decisions made at the level of individual modalities;
- The individual modalities are assumed to be statistically independent given the class;
- In the late fusion case, there are changes from a decision theoretic point of view.

d. CMU

i. https://www.youtube.com/watch?v=2xr4P0WGKSA

ii. 


iii. Train classifier first

# Lecture 20: Classifier Combination

12 May 2022 16:18

## Late fusion

1.

# Lab2

07 December 2021 11:28

## 指标计算

1. https://blog.csdn.net/program_developer/article/details/79937291?ops_request_misc=%257B%2522request%255Fid%2522%253A%2522163884763616780274167581%2522%252C%2522scm%2522%253A%252220140713.130102334..%2522%257D&request_id=163884763616780274167581&biz_id=0&utm_medium=distribute.pc_search_result.none-task-blog-2~all~baidu_landing_v2~default-4-79937291.first_rank_v2_pc_rank_v29&utm_term=%E9%94%99%E8%AF%AF%E7%8E%87&spm=1018.2226.3001.4187
2. GDA
    a. https://blog.csdn.net/qq_30091945/article/details/81508055
    b. https://linkin1005.blog.csdn.net/article/details/39054023?spm=1001.2101.3001.6650.3&utm_medium=distribute.pc_relevant.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-3.no_search_link&depth_1-utm_source=distribute.pc_relevant.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-3.no_search_link
    c. https://www.cnblogs.com/jcchen1987/p/4424436.html
3. Error theory
    a. https://linkin1005.blog.csdn.net/article/details/42563229
4. 
    a. https://www.geeksforgeeks.org/gaussian-discriminant-analysis/

## 先验概率和后验概率

1. 先验概率: 专家依据经验得出的结论
2. 后验概率: 通过贝叶斯定理, 用先验概率和似然函数求得

## 贝叶斯分类器

1. https://www.cnblogs.com/listenfwind/p/10040624.html
2. https://blog.csdn.net/qq_25948717/article/details/81744277
3. https://blog.csdn.net/weixin_44049128/article/details/86511224

# Exam

30 November 2021 19:33

| | 2021 | 2020 | 2019 | mock |
|---|---|---|---|---|
| Contingency table | 1-1 | | | |
| Chi square formula | 1-2 | | | |
| Provide Np | 1-3 | | | |
| expectation values | 1-4 | | | |
| Dof of chi square | 1-5 | | | |
| The value of chi square | 1-6 | | | |
| Whether np can be rejected | 1-7 | | | |
| Prior rule formula | 2-1 | | | |
| Prior rule class | 2-2 | | | |
| Posterior formula | 2-3 | | | |
| Posterior rule class | 2-4 | | | |
| Essay | 3 | 3 | 3 | |
| Signal processing | | 1-1 | | |
| Bayes decision rule, Excepted loss/Conditional risk | | 1-2 | | |
| Facial expressions | | 2-1 | | |
| Multimodality in communication and life science | | 2-2 | | |
| Early fusion/ late fusion | | | 1-1 | |
| Combination classifiers | | | 1-2 | |
| Judgment studies and reliability | | | 2-1 | |
| Encoders/decoders | | | 2-2 | |
| correlation | | | | 1 |
| covariance | | | | 2 |
| Statistically relationship | | | | 3 |
| Statistically difference | | | | 4 |
| | | | | |
| | | | | |

## 2022: to the exam

1. Mice and learning: 149
2. Exam
    a. 相关性
    b. 协方差, 公式
    c. 统计学差异: 先求r, 转换成t, t有阈值

## 2019-1.1: early and late fusion

1. Early fusion
    a. The early fusion is the concatenation of the feature vectors extracted from the data captured through multiple sensors
    b. The concatenation can be treated like any other vector
    c. In the early fusion case, there are no changes from a decision theoretic point of view
    d. Formula

2. Late fusion
    a. The late fusion is the combination of decisions made at the level of individual modalities
    b. The individual modalities are assumed to be statistically independent given the class
    c. In the late fusion case, there are changes from a decision theoretic point of view
    d. Formula

## 2019-1.2: combination rules

1. Max rule
2. Min rule

## 2019-2.1

1. Reliability

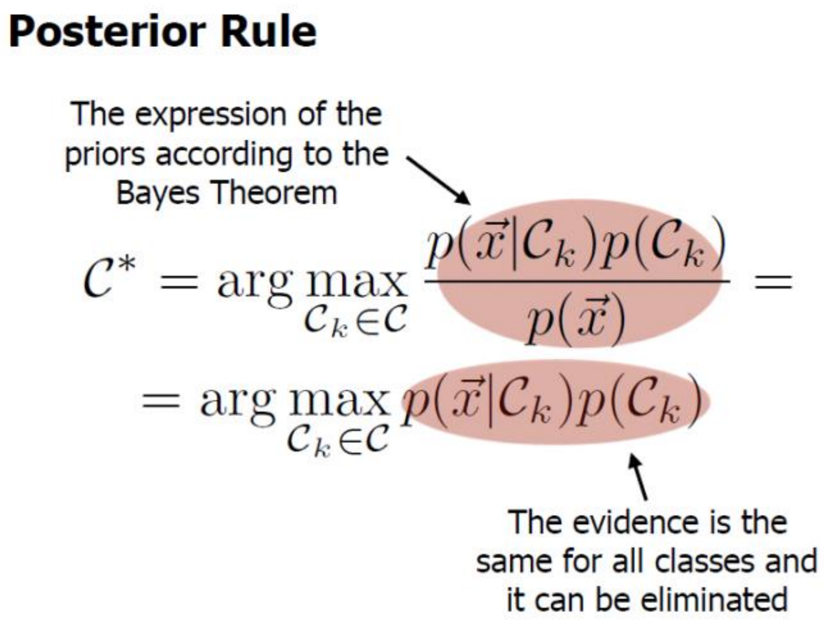
## 2019-2.2

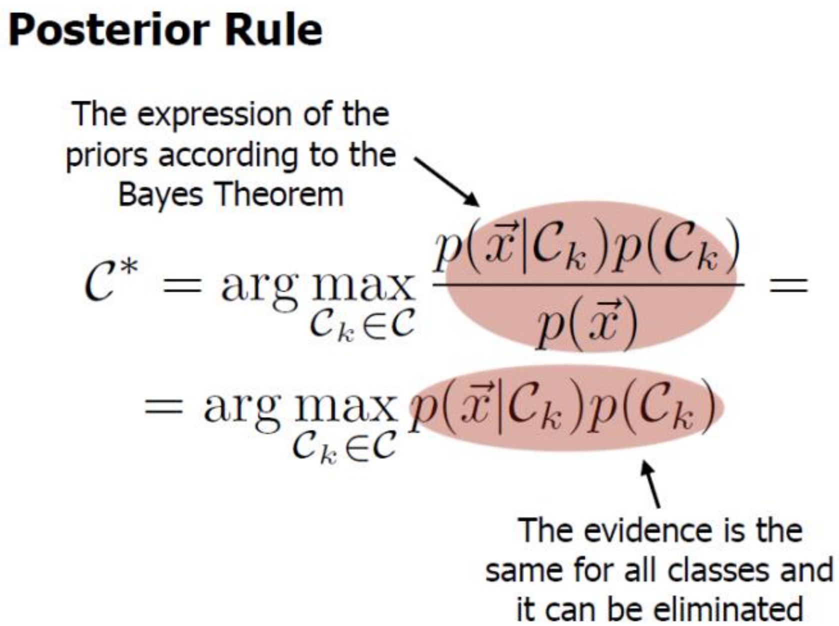
1. Encoder
2. Decoder

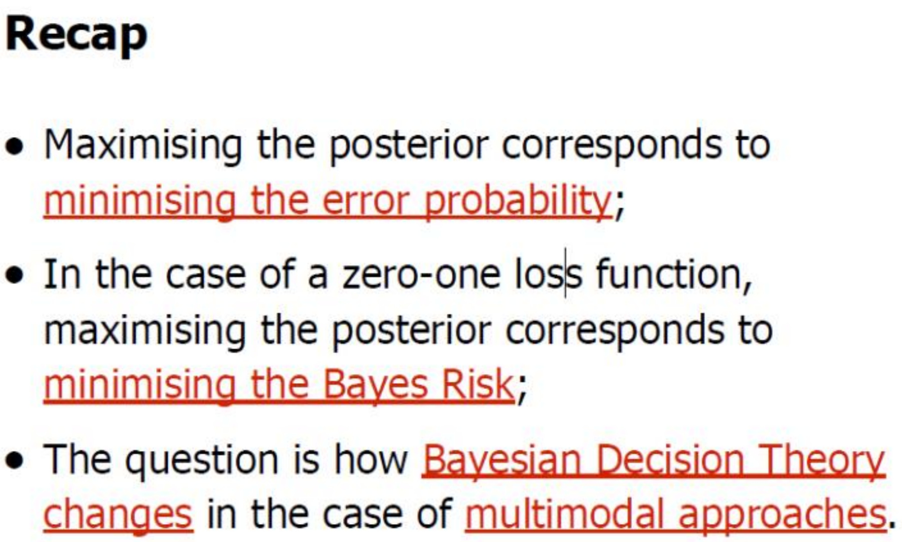
## 2020-1.1

## 2021-2-3: posterior rule

1. P715

a. 


b. 


c. 


## 2021-2

1. Prior rule
    a. Image you are having dinner in a steak house. There are M steak dishes the

steak house serve, but your money only can afford one dish.

   b. The decision you make is Ci, i stands for each dish. The probability of picking each dish is p(Ci), named a-priori probability. The one dish with the highest a-priori probability is the final decision C* you make.

2. Posterior rule
   a. Image you are in the same condition above. You decide to have steak(prior), but you do not know the degree of the steak: rare, medium rare, medium, medium well, or well done. So you ask the waiter for advice. He tells you the percentage of people for each degree at each dish(likelihoods). According to this, you will product them together(posterior) and pick the one with the highest score(C*).

## 3-essay

1. Establish connection between experimental study and methodologies
2. Research hypothesis and results
3. Psychometric instruments

# 下一步复习计划

1. 2021期末考试题
2. Case study
3. 准备最后一题
4. gg

# Introduction

08 November 2022 16:14

For more information, go to
https://github.com/Zak3225/Glasgow_COMPSCI4080_CSI