

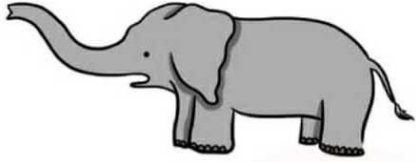
# Statistics

*for* Data Science

گروه دایچه . dayche.com



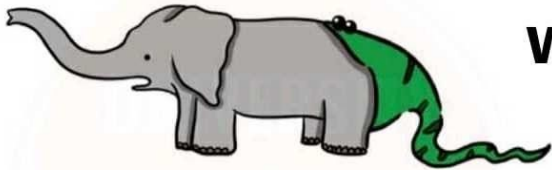
**Statistics**



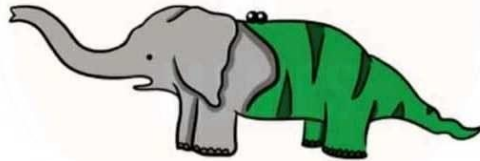
**Computer Science**



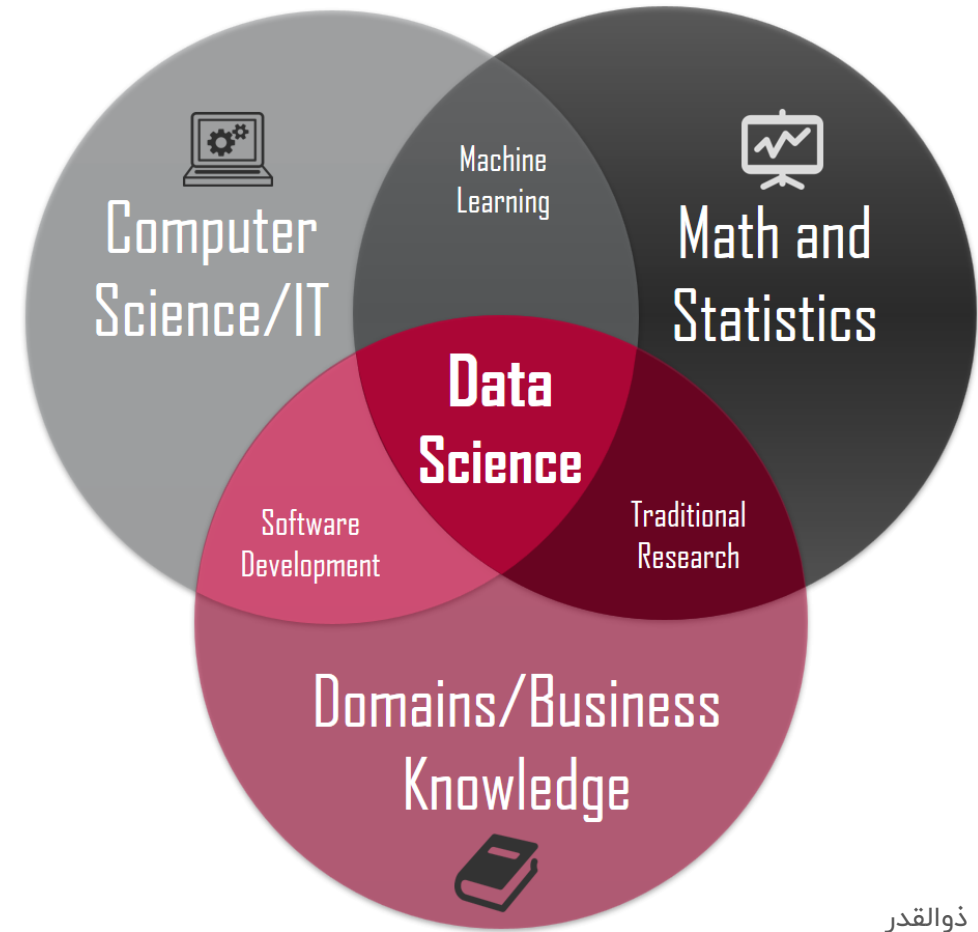
**We will work together**



**Please teach me Statistics**



**Now I am a Data Scientist**



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

گروه دایچه | dayche.com




مبانی احتمال

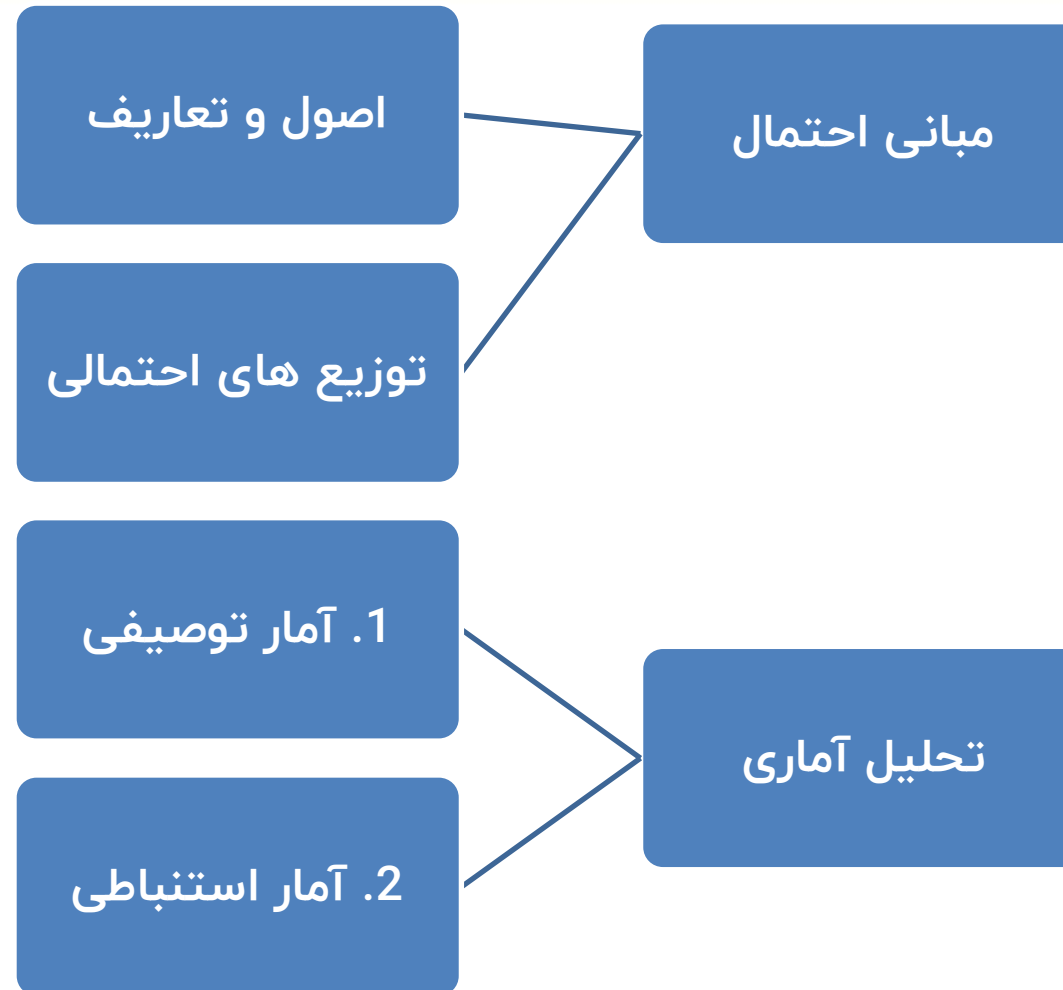
تحلیل آماری

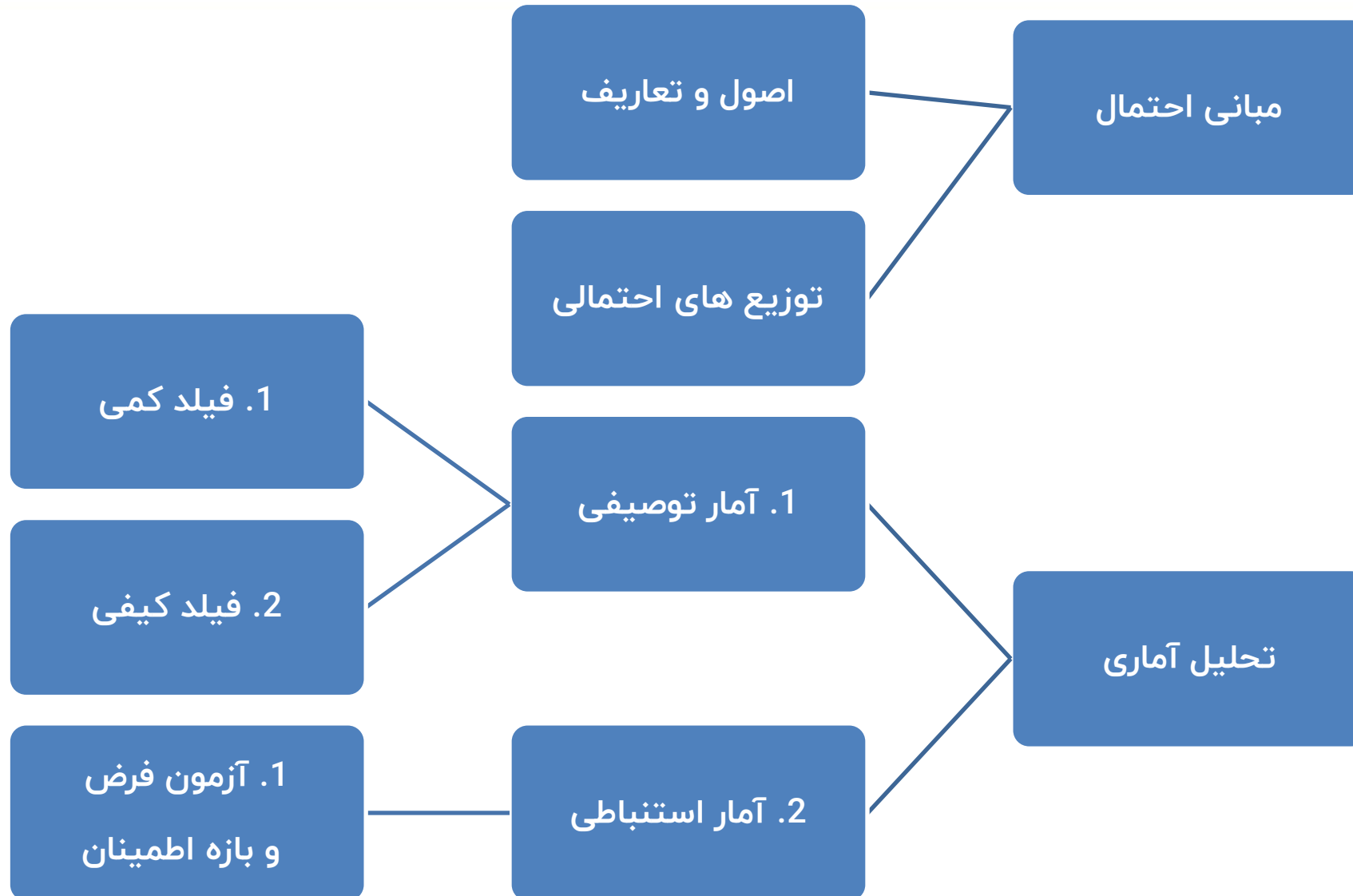
تولید محتوا: زهرا ذوالقدر

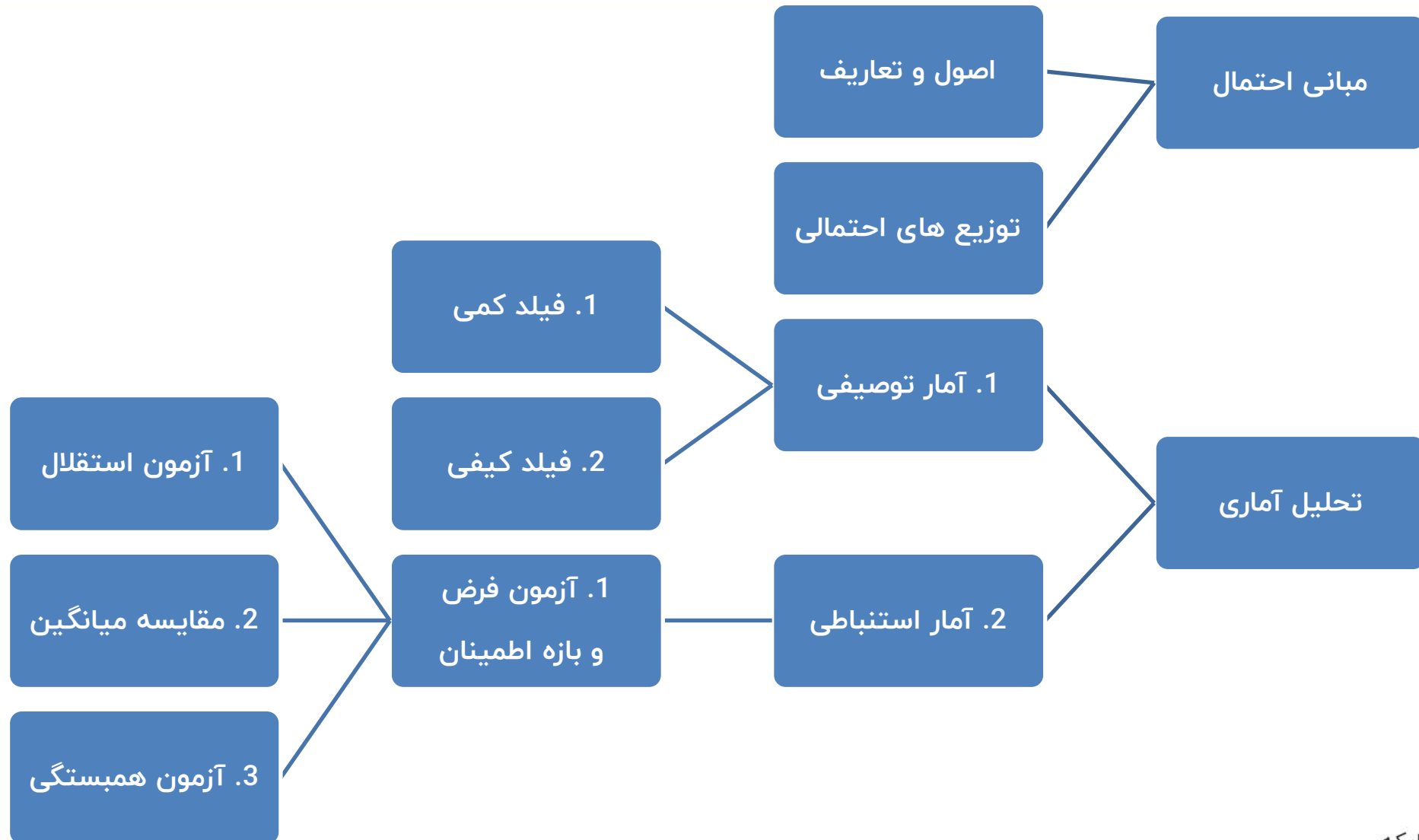
daychegroup 

daychegroup 

dayche.com | گروه دایکه 







# Probability Theory

نئوری احتمال

گروه دایچه . dayche.com



# نئوری احتمال

## مفاهیم پایه

### آمار و احتمال


- احتمال، شاخه ای است از ریاضیات جهت مطالعه فضای عدم قطعیت.
- آمار، رشته ای است بر پایه نظریه احتمالات به منظور جمع آوری، تجزیه و تحلیل، تفسیر و ارائه اطلاعات مفید از داده ها.



تولید محتوا: زهرا ذوالقدر

daychegroup 

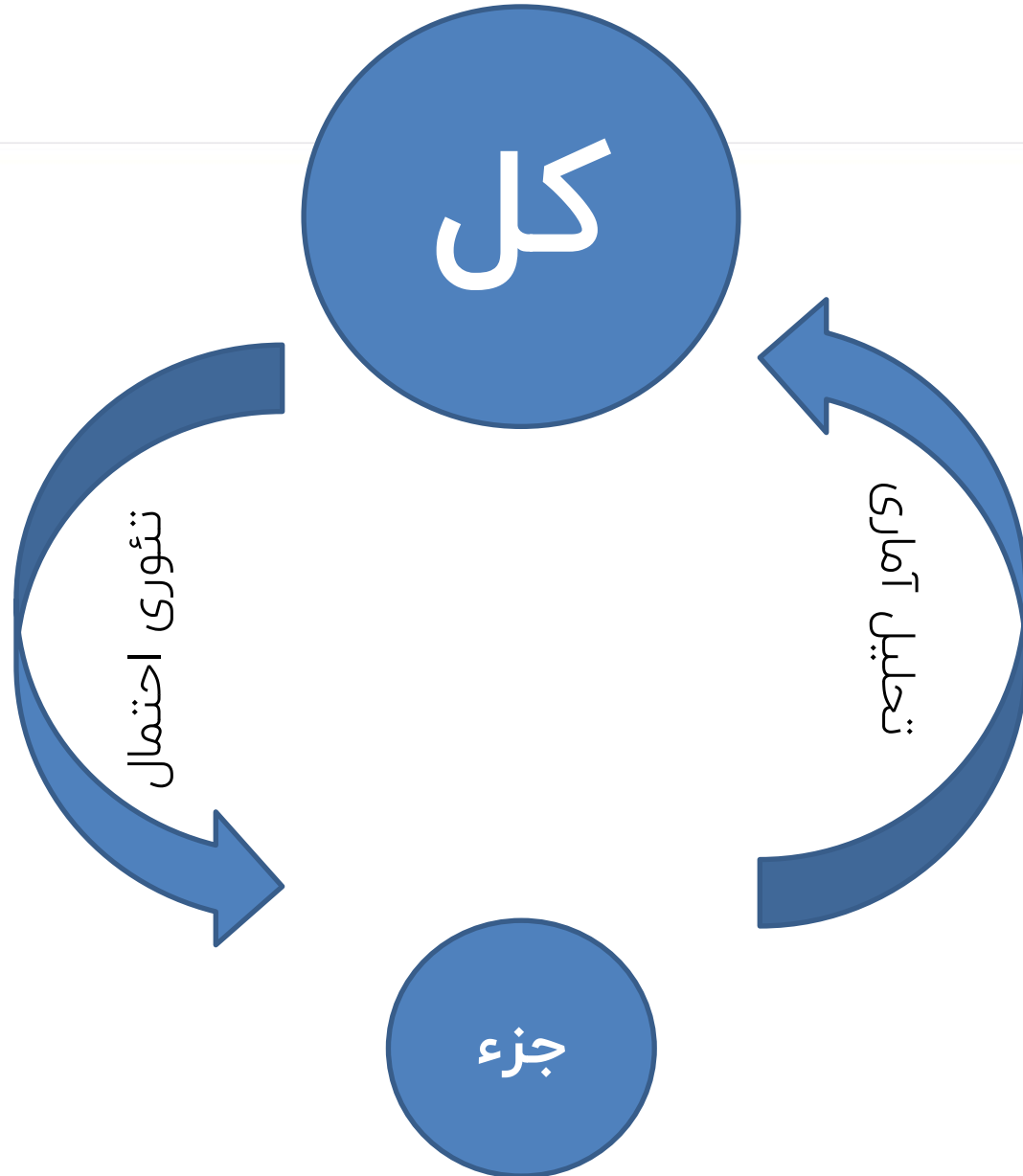
daychegroup 

dayche.com | گروه دایچه 



# تئوری احتمال

## مفاهیم پایه




□ تئوری احتمال: تصمیم گیری از کل به جزء

□ تحلیل آماری: تصمیم گیری از جزء به کل

تولید محتوا: زهرا ذوالقدر

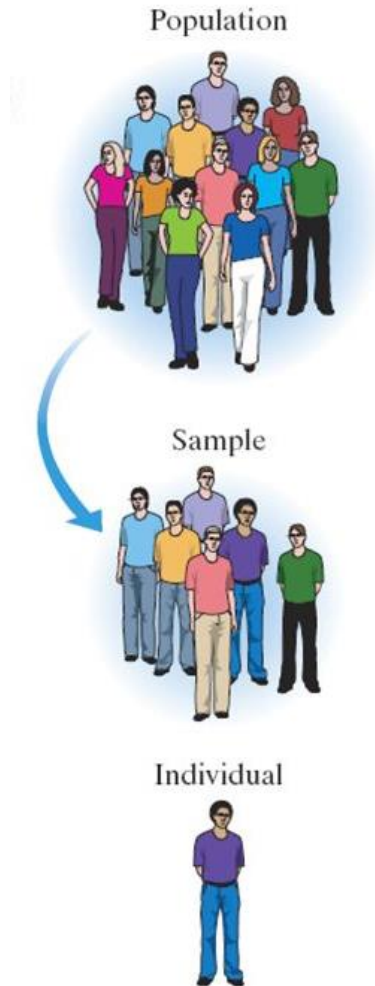
daychegroup 

daychegroup 

dayche.com | گروه دایکه 

# تئوری احتمال

## مفاهیم پایه

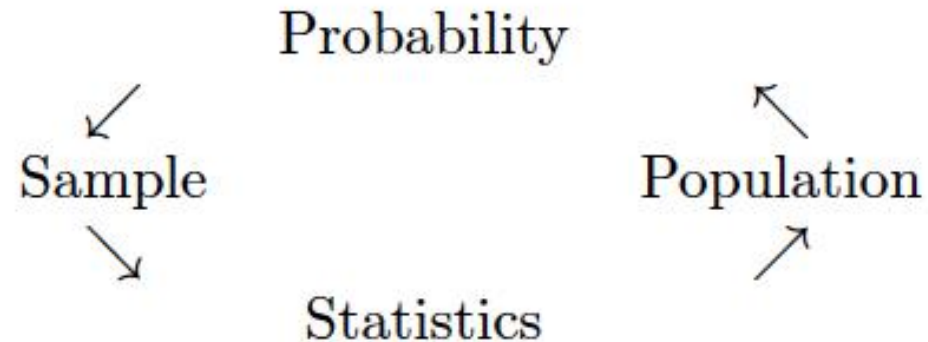


□ جامعه آماری (POPULATION):

گروهی از افراد یا اشیاء که در ویژگی یا ویژگی‌های مورد پژوهش مشترک بوده و با هدف و موضوع تحقیق مرتبط باشد.

□ نمونه (SAMPLE):

بخشی از جامعه آماری که مورد مشاهده قرار گرفته است.



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه

# تئوری احتمال

## مفاهیم پایه

### □ آزمایش تصادفی: (Random Experiment)

آزمایشی که نتایج آن با اطمینان کامل قابل پیش بینی نباشد. مثل:  
پرتاب یک سکه؛

افتتاح حساب توسط مشتری بانک؛

گرفتن تاکسی با استفاده از اپلیکیشن؛


ابتلای یک فرد به بیماری Covid19؛



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 



### فضای نمونه (Sample Space) □

مجموعه تمام نتایج ممکن یک آزمایش تصادفی

$$S=\{H,T\}$$

پرتاب یک سکه

$$S=\{\text{قرض الحسنه، جاری، کوتاه مدت، بلند مدت}\}$$

افتتاح حساب توسط مشتری بانک

$$S=\{\text{زمان انتظار تا 5 دقیقه}\}$$

گرفتن تاکسی با استفاده از اپلیکیشن

$$S=\{\text{بهبودی، فوت}\}$$

ابتلای یک فرد به بیماری Covid19



### پیشامد (Event) □

هریک از زیر مجموعه های فضای نمونه

$A=\{H\}$      $B=\{T\}$      $C=\{H,T\}$      $D=\{ \}$

$A=\{\text{جاری}\}$      $B=\{\text{بلند مدت}\}$      $C=\{\text{قرض الحسنه، کوتاه مدت}\}$     ...     $D=\{ \}$

$A=\{\text{بیشتر از 3 دقیقه}\}$      $B=\{\text{کمتر از 2 دقیقه}\}$      $C=\{\text{بین 2 تا 5 دقیقه}\}$     ...     $D=\{ \}$

$A=\{\text{فوت}\}$         $B=\{\text{بهبودی}\}$         $C=\{\text{فوت، بهبودی}\}$     ...     $D=\{ \}$

پرتاب یک سکه

افتتاح حساب توسط مشتری بانک

گرفتن تاکسی با استفاده از اپلیکیشن

ابتلای یک فرد به یک بیماری

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه

### احتمال (Probability) □

شاخه ای از ریاضیات است که به صورت عددی باورپذیری رخداد یک پیشامد را بیان می کند.

$$P(A)=P\{H\}=1/2 \quad B=\{T\} \quad C=\{H,T\} \quad D=\{ \}$$

پرتاب یک سکه

$$A=\{\text{جاری}\} \quad B=\{\text{بلند مدت}\} \quad C=\{\text{قرض الحسنه، کوتاه مدت}\} \quad \dots \quad P(D)=P(\{\})=0$$

افتتاح حساب توسط مشتری

$$A=\{\text{بیشتر از 3 دقیقه}\} \quad P(B)=P\{\text{کمتر از 2 دقیقه}\}=2/5 \quad C=\{\text{بین 2 تا 5 دقیقه}\} \quad \dots \quad D=\{ \}$$

گرفتن تاکسی با استفاده از اپلیکیشن

$$A=\{\text{فوت}\} \quad B=\{\text{بهبودی}\} \quad P(C)=P\{\text{فوت، بهبودی}\}=1 \quad \dots \quad D=\{ \}$$

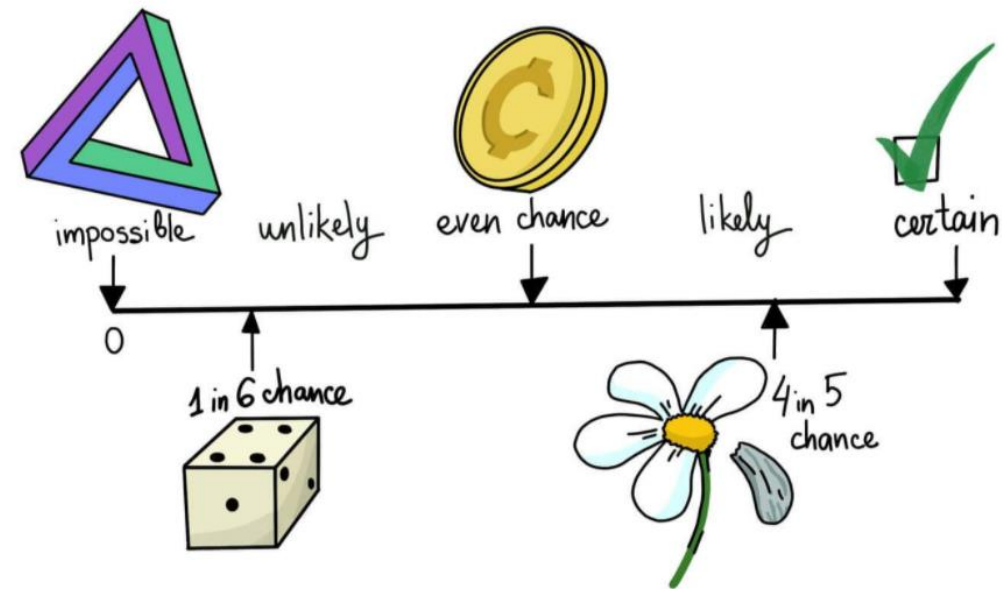
ابتلای یک فرد به یک بیماری Covid19

# تئوری احتمال

## مفاهیم پایه

### □ احتمال (Probability)

شاخه ای از ریاضیات است که به صورت عددی باورپذیری رخداد یک پیشامد را بیان می کند.



تولید محتوا: زهرا ذوالقدر

daychegroup

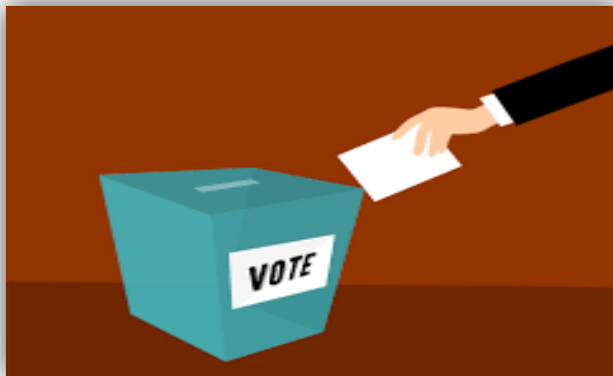
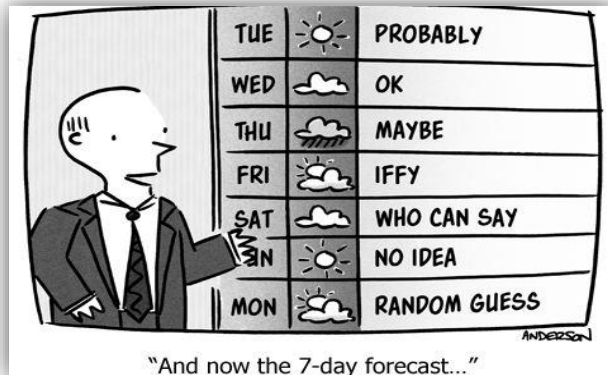
daychegroup

dayche.com | گروه دایچه

# تئوری احتمال

مفاهیم پایه

□ احتمال در زندگی روزمره



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

گروه دایچه | dayche.com



# نئوری احتمال

## مفاهیم پایه

### انواع تعابیر احتمال □

- به طریق شخصی  
بر پایه قضاوت و داوری بر اساس **اطلاعات و دانسته های قبلی** فرد.
- به طریق کلاسیک  
بر پایه مفهوم **هم شانس بودن** پیشامدهای ساده.
- به طریق فراوانی نسبی  
بر پایه **تقریبی از فراوانی نسبی** رخداد پیشامدهای مورد نظر، در **تکرارهای زیاد** آزمایش تحت **شرایط یکسان**.

تولید محتوا: زهرا ذوالقدر

daychegroup 

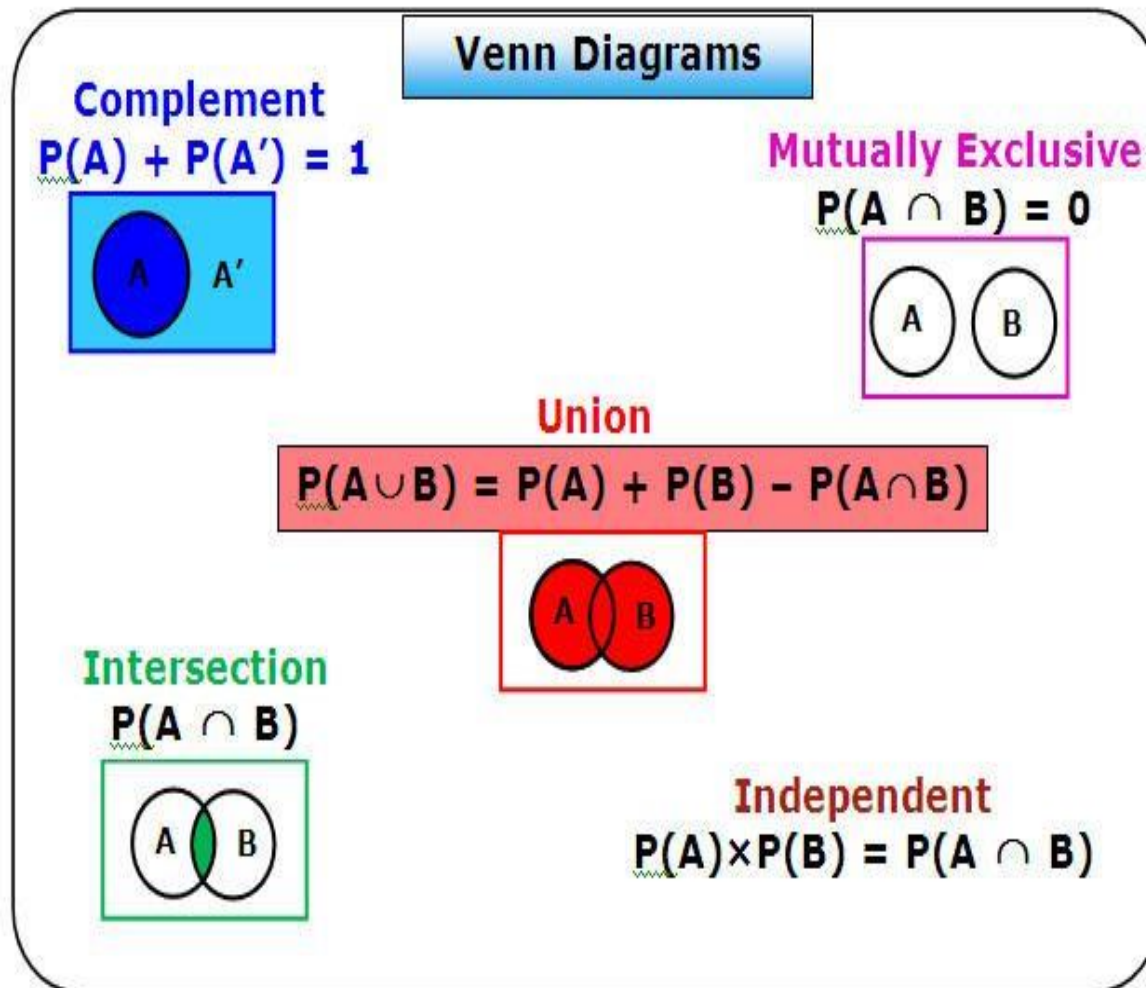
daychegroup 

dayche.com | گروه دایکه 

### □ احتمال در نمودار Venn

توسعه روابط احتمالی بر اساس نظریه مجموعه ها:


- مکمل
- اجتماع
- اشتراک
- ناسازگار
- استقلال



تولید محتوا: زهرا ذوالقدر

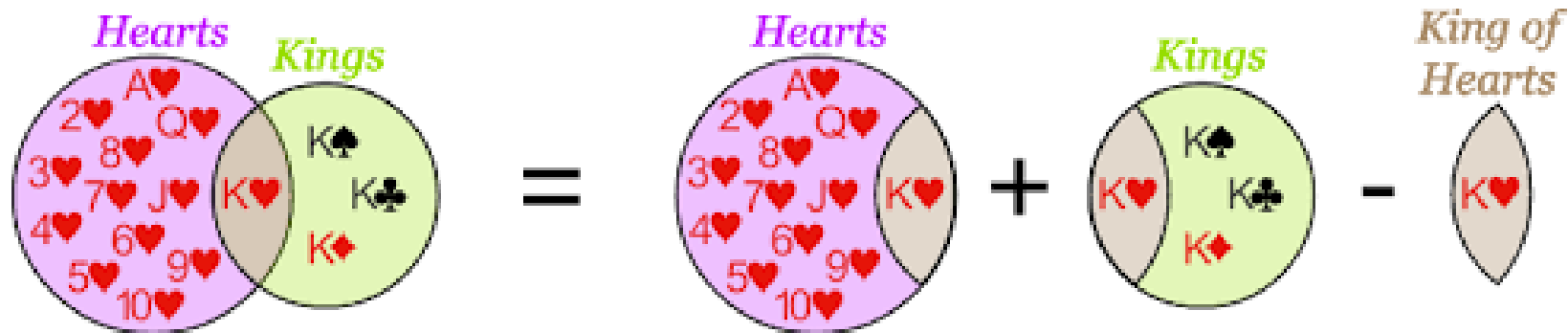
daychegroup 

daychegroup 

dayche.com | گروه دایکه 

احتمال در نمودار Venn □

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



# نئوری احتمال

## مفاهیم پایه

### پیشامد های ناسازگار و قانون جمع احتمال

(Mutually Exclusive Events & the Addition Rule)

○ پیشامد هایی که امکان وقوع همزمان ندارند، مثل:

داشتن گروه خونی 0 یا A

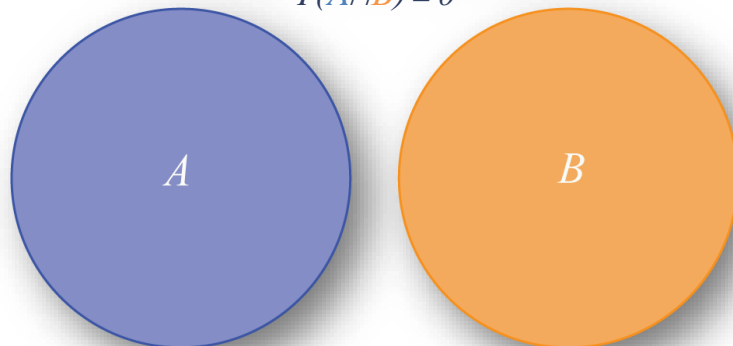
دویدن به جلو و دویدن به عقب در یک زمان

تماشا کردن تلویزیون و نداشتن تلویزیون

○ در این حالت احتمال وقوع یکی از دو پیشامد، برابر مجموع احتمالات آنهاست:

$$P(A \cup B) = P(A) + P(B)$$

$$P(A \cap B) = 0$$




Note: No overlap between A and B

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

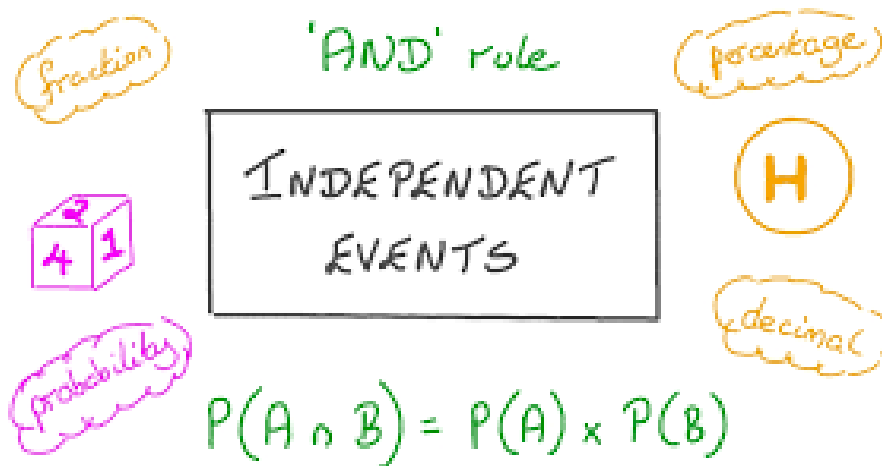
dayche.com | گروه دایچه 

### پیشامد های مستقل و قانون ضرب احتمال

(Independent Events & the Multiplication Rule)

- دو پیشامدی که وقوع یا عدم وقوع یکی از آنها، تاثیری در وقوع یا عدم وقوع دیگری ندارد، مثل:
  - مرد بودن و داشتن گروه خونی O
  - با تاکسی به خانه رفتن و آمدن مهمان سرزده
  - بارندگی هوا و برنده شدن در لاتاری

- در این حالت احتمال توام دو پیشامد، برابر حاصل ضرب احتمالات آنهاست:



پیشامد های ناسازگار / مستقل / وابسته □

Mutually  
Exclusive




Independent Dependent



تولید محتوا: زهرا ذوالقدر

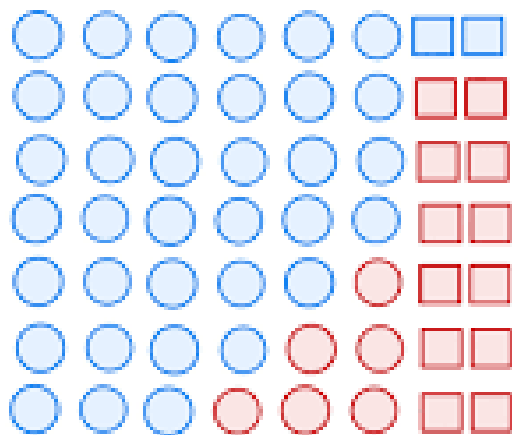
daychegroup 

daychegroup 

dayche.com | گروه دایچه 

### احتمال شرطی (Conditional Probability) □

در بسیاری از پدیده های تصادفی، اغلب آگاهی از رخداد پیشامد B در محاسبه احتمال پیشامد A تاثیر می گذارد.



#### Conditional Probability Formula

$$P(A | B) = \frac{\text{Probability of } A \text{ and } B}{\text{Probability of } B}$$

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

$$P(\text{circle} | \text{red}) = \frac{P(\text{circle} \cap \text{red})}{P(\text{red})} = \frac{5}{10}$$

### احتمال شرطی (Conditional Probability) □

در بسیاری از پدیده های تصادفی، اغلب آگاهی از رخداد پیشامد B در محاسبه احتمال پیشامد A تاثیر می گذارد.

#### Conditional Probability Formula

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Probability of A and B

Probability of A given B

Probability of B

	Number of times students visited tutoring			Total
	One or fewer times	Two to three times	Four or more times	
Full time student	12	25	8	45
Part time student	2	5	6	13
Total	14	30	14	58


$$P(\text{four or more times} | \text{full time student}) = \frac{8}{45} \approx 0.18$$

find given

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

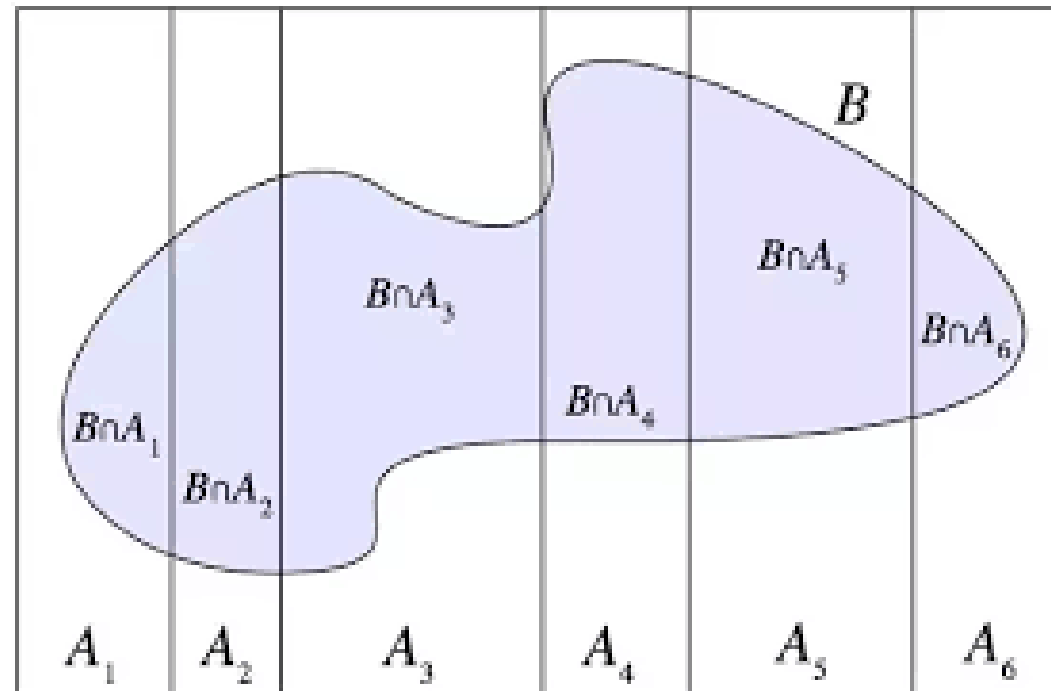


قانون احتمال کل (Law of total probability) □

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) + P(B \cap A_4) + P(B \cap A_5) + P(B \cap A_6)$$

$$= P(B | A_1) P(A_1) + P(B | A_2) P(A_2) + P(B | A_3) P(A_3) + P(B | A_4) P(A_4) + P(B | A_5) P(A_5) + P(B | A_6) P(A_6)$$

$$= \sum_i P(B | A_i) P(A_i)$$



# تئوری احتمال

## مفاهیم پایه

قضیه بیز (bayes theorem) □

ایده: روشی برای دسته بندی پدیده ها بر اساس احتمال وقوع یا عدم وقوع یک پدیده.

Number of occurrences	Beard: No beard:		sum
	B	$\bar{B}$	
Astigmatic: A	2	3	5
Not astigmatic: $\bar{A}$	6	9	15
sum	8	12	20

B	$\bar{B}$
A	3

B	$\bar{B}$
A	3
$\bar{A}$	9

B	$\bar{B}$
A	3
$\bar{A}$	9

$$P(B, \text{ given } A) \cdot P(A) = P(B|A) \cdot P(A)$$

$$\frac{2}{2+3} \cdot \frac{2+3}{2+3+6+9} = \frac{2}{2+3+6+9}$$

B	$\bar{B}$
A	6

B	$\bar{B}$
A	3
$\bar{A}$	9

B	$\bar{B}$
A	3
$\bar{A}$	9

$$P(A, \text{ given } B) \cdot P(B) = P(A|B) \cdot P(B)$$

$$\frac{2}{2+6} \cdot \frac{2+6}{2+3+6+9} = \frac{2}{2+3+6+9}$$

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$\therefore P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایچه



### قضیه بیز (bayes theorem) □

بسط قضیه بیز بر اساس قانون احتمال کل


$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

### قضیه بیز (bayes theorem) □

مثال:

با فرض اینکه همه بیماران مبتلا به سرطان پانکراس دارای نشانه‌ها و علائم خاصی باشند، این موضوع دلیلی قطعی بر وجود سرطان در افرادی که دارای آن نشانه‌های خاص باشند نیست.


فرض کنید احتمال بروز سرطان پانکراس  $1/100000$  باشد و میدانیم  $10/100000$  افراد، سالم هستند و دارای نشانه‌ها و علائم خاص این بیماری نیز می‌باشند.

با این شرایط احتمال مبتلا بودن یک فرد به سرطان پانکراس به شرط اینکه دارای علائم باشد چه میزانی است؟

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

قضیه بیز (bayes theorem) □

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\neg A)P(\neg A)}$$

Cancer \ Symptom	Yes	No	Total
Yes	1	10	11
No	0	99989	99989
Total	1	99999	100000


ادامه مثال:

$$\begin{aligned} P(\text{Cancer}|\text{Symptoms}) &= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms})} \\ &= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer}) + P(\text{Symptoms}|\text{Non-Cancer})P(\text{Non-Cancer})} \\ &= \frac{1 \times 0.00001}{1 \times 0.00001 + (10/99999) \times 0.99999} = \frac{1}{11} \approx 9.1\% \end{aligned}$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# نئوری احتمال

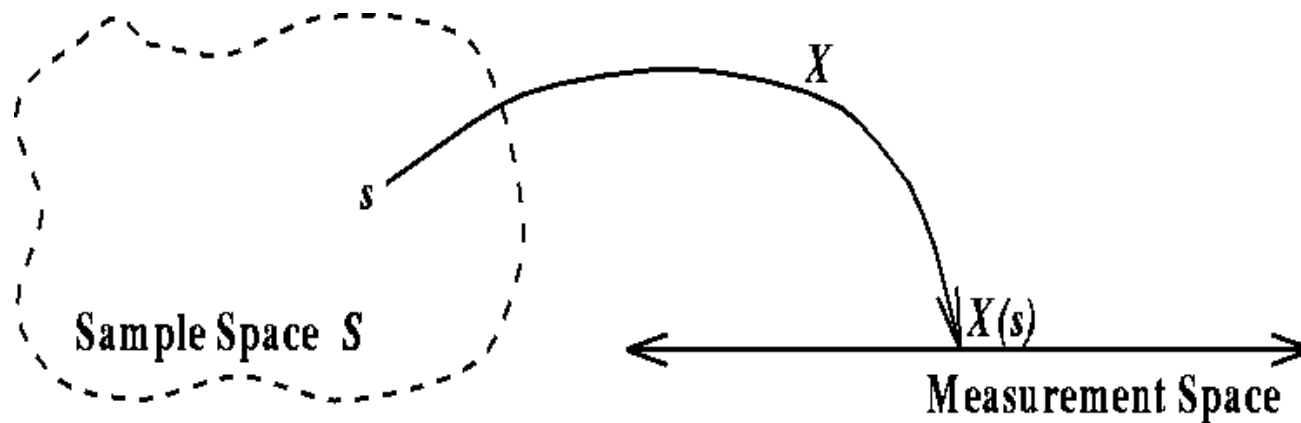
## متغیر تصادفی

### متغیر تصادفی □

یک تابع حقیقی روی فضای نمونه است. مثل:

تعداد شیرها در پرتاب ده سکه


رتبه اولین کلیک در نتایج صفحه اول گوگل



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

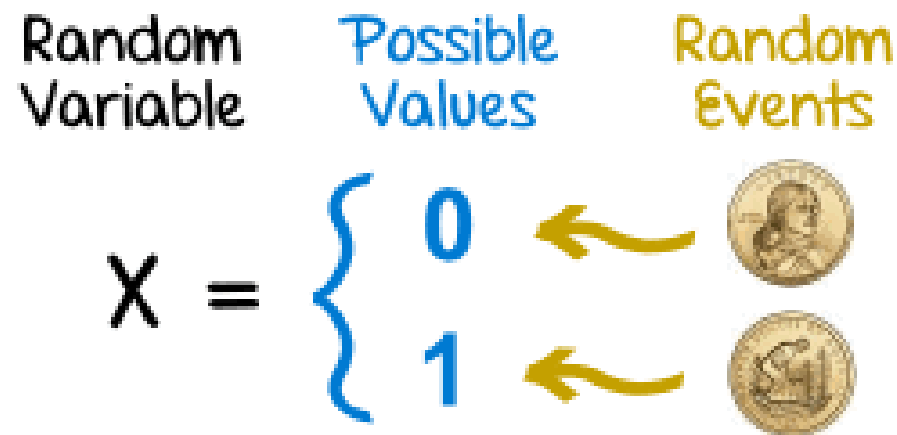
# نئوری احتمال

## متغیر تصادفی

### □ متغیر تصادفی گسسته:

برد یک متغیر تصادفی گسسته (کلیه مقادیر امکانپذیر)، یک مجموعه متناهی یا نامتناهی شمارش پذیر از اعداد حقیقی است. به فرم بسته ای که مقدار احتمال هر پیشامد را نشان می دهد **تابع جرم احتمال** یا بطور خلاصه **تابع احتمال** گویند.


مثال: حالت ساده پرتاب یک سکه:



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# نئوری احتمال

## متغیر تصادفی



مثال: یک سکه سالم را دوبار پرت می کنیم.  
فضای نمونه:

$$S = \{HH, TT, HT, TH\}$$

$X$ : متغیر تصادفی تعداد شیرها

$$P(X=0) = P(\{TT\}) = \frac{1}{4}$$

$$P(X=2) = P(\{HH\}) = \frac{1}{4}$$

$$P(X=1) = P(\{HT, TH\}) = \frac{1}{2}$$

$$\text{تابع احتمال } f(x) = \begin{cases} \frac{1}{4} & \text{if } x = 0 \\ \frac{1}{2} & \text{if } x = 1 \\ \frac{1}{4} & \text{if } x = 2 \end{cases}$$

تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

dayche.com | گروه دایکه



### متغیر تصادفی پیوسته: □

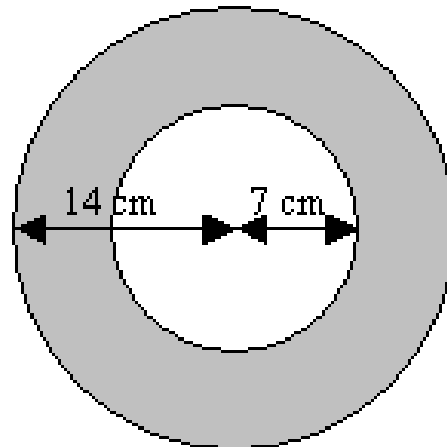
برد یک متغیر تصادفی پیوسته (کلیه مقادیر امکانپذیر)، یک مجموعه نامتناهی ناشمارا از اعداد حقیقی است. به فرم بسته ای که چگالی احتمال هر پیشامد را نشان می دهد **تابع چگالی** گویند.

مثال: یک نقطه به تصادف از دایره ای به مرکز 0 و شعاع  $r$  انتخاب می کنیم.

$X$ : متغیر تصادفی طول نقطه انتخابی از مرکز دایره

$$f(x) = \frac{2x}{r^2}$$

$$0 \leq x \leq r$$



# تئوری احتمال

## متغیر تصادفی




### جمع بندی

- متغیر های تصادفی یک تابع حقیقی هستند.
- متغیر های تصادفی، احتمال نیستند.
- بسته به برد (فضای امکانپذیر) متغیر های تصادفی به دو دسته گسسته و پیوسته تقسیم می شوند.
- به فرم بسته مقدار یا چگالی احتمال متغیرهای تصادفی گسسته و پیوسته به ترتیب تابع جرم احتمال و تابع چگالی می گویند.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

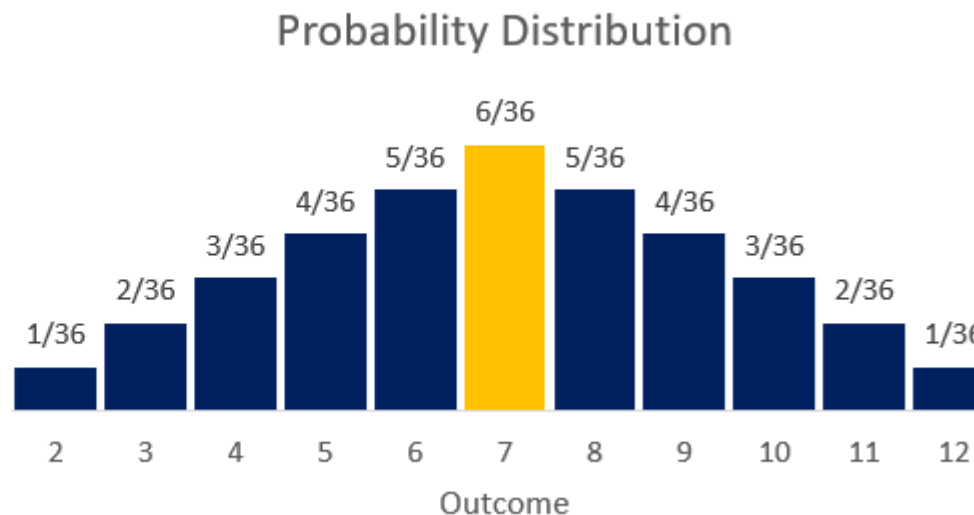
dayche.com | گروه دایچه 

# تئوری احتمال

## توزیع های احتمالی

- برخی از متغیرهای تصادفی از تابع احتمال یا تابع چگالی خاصی پیروی می کنند که به توزیع های احتمالی استاندارد شناخته می شوند.
- شناخت و درک درست از ساختار مدل احتمالی این توزیع ها، منجر به درک درست و دقیق از رفتار متغیر تصادفی در حل مسائل خواهد شد.


Outcomes	Occurrence	Probability
2	1	1/36
3	2	2/36
4	3	3/36
5	4	4/36
6	5	5/36
7	6	6/36
8	5	5/36
9	4	4/36
10	3	3/36
11	2	2/36
12	1	1/36



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تئوری احتمال

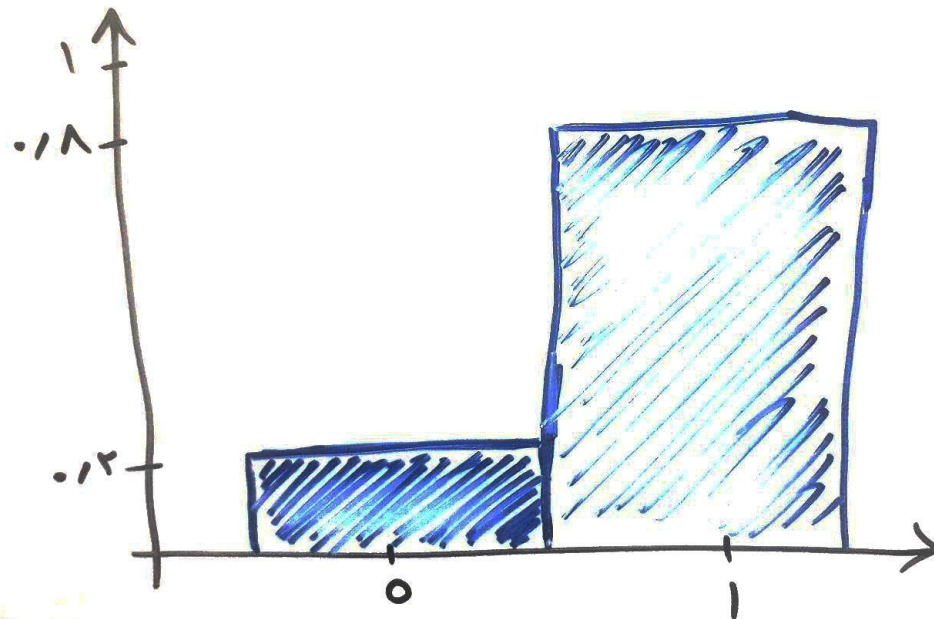
## توزیع های احتمالی

### توزیع برنولی □

آزمایشی داریم که دارای دو برآمد (شکست یا پیروزی) است، اگر  $X$  نمایانگر تعداد پیروزی ها باشد،  $X$  دارای توزیع برنولی با پارامتر  $p$  (احتمال پیروزی) است.

$$p = 0.18$$

تابع احتمال:



$$P(X = x) = p^x (1 - p)^{1-x}$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

### □ توزیع دو جمله ای

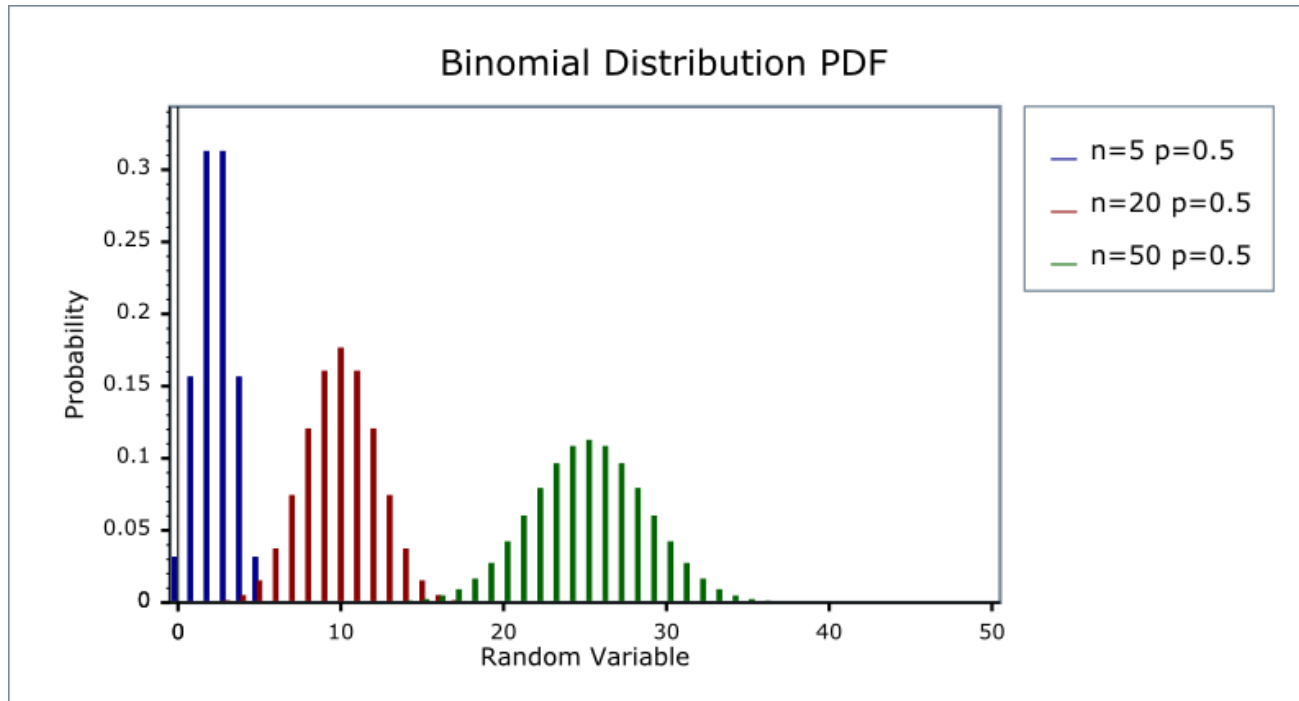
اگر یک آزمایش با دو برآمد (شکست یا پیروزی) را  $n$  بار تکرار کنیم و  $X$  نمایانگر تعداد پیروزی ها باشد،  $X$  دارای توزیع دو جمله ای با پارامترهای  $n$  و  $p$  (احتمال پیروزی) است.

تابع احتمال:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

○ توزیع دو جمله ای از حاصل جمع  $n$  توزیع برنولی به دست می آید. در واقع توزیع برنولی حالت خاصی از توزیع دو جمله ای محسوب می شود که در آن  $n$  برابر با یک باشد.

### توزیع دو جمله ای □



مثال: سه توزیع دو جمله ای با تعداد  
آزمایش های متفاوت


$$\mu_X = np$$

$$\sigma_X^2 = np(1-p)$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

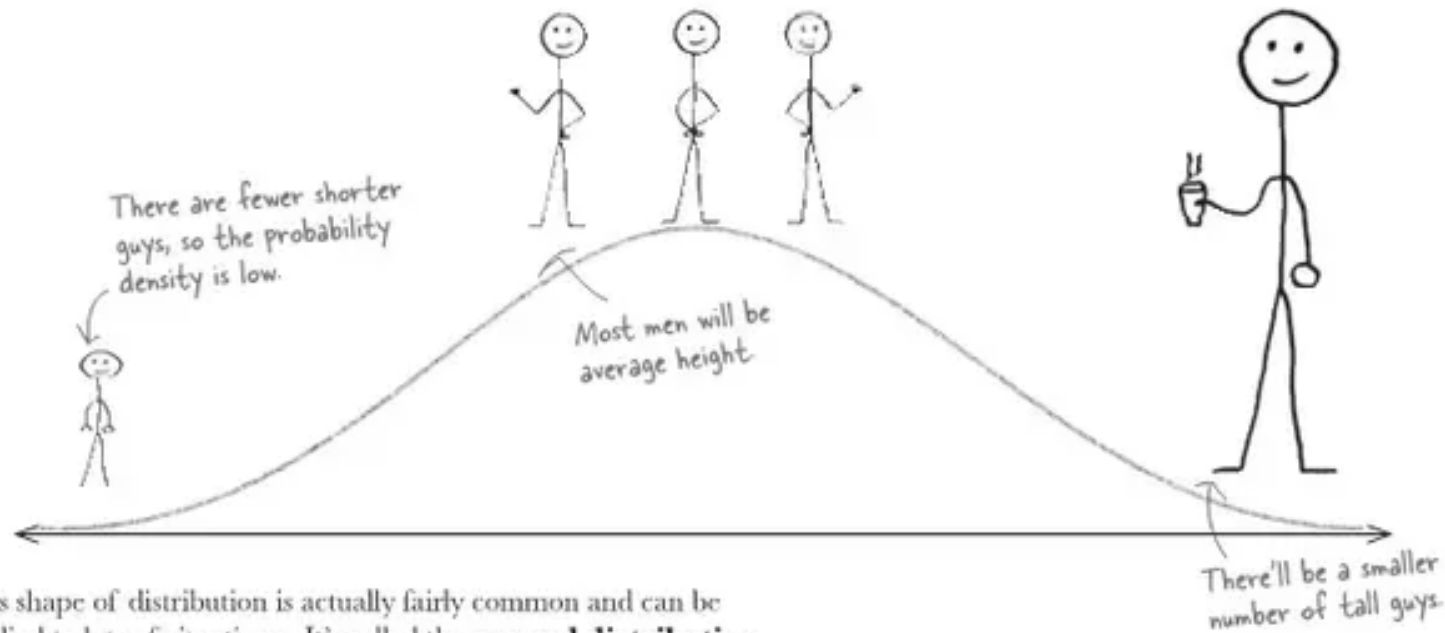
daychegroup 

dayche.com | گروه دایکه 

# تئوری احتمال

## توزیع های احتمالی

### توزیع نرمال □



This shape of distribution is actually fairly common and can be applied to lots of situations. It's called the **normal distribution**.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

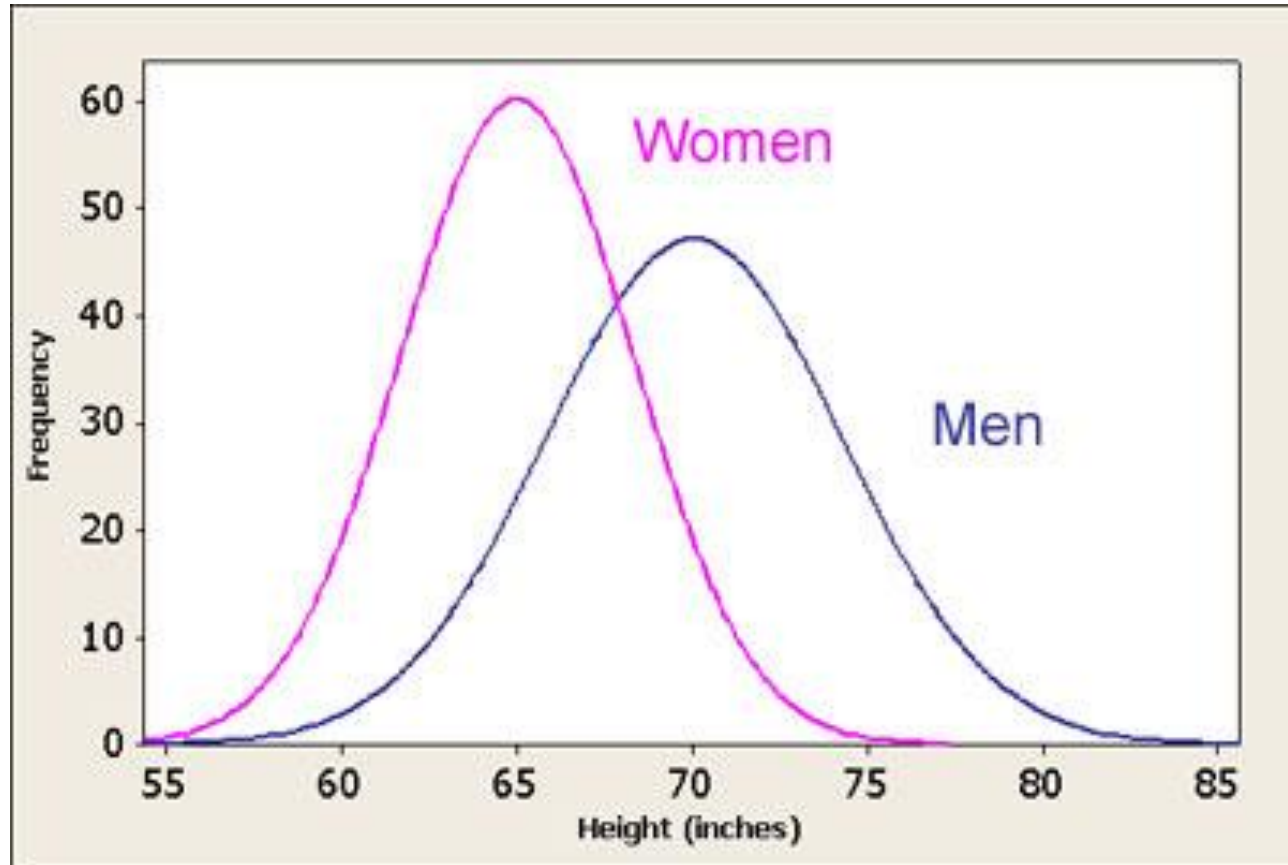
گروه دایچه | dayche.com 

# تئوری احتمال

توزیع های احتمالی

توزیع نرمال □


مثال: توزیع قد افراد در زنان و مردان



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 



# نئوری احتمال

## توزیع های احتمالی

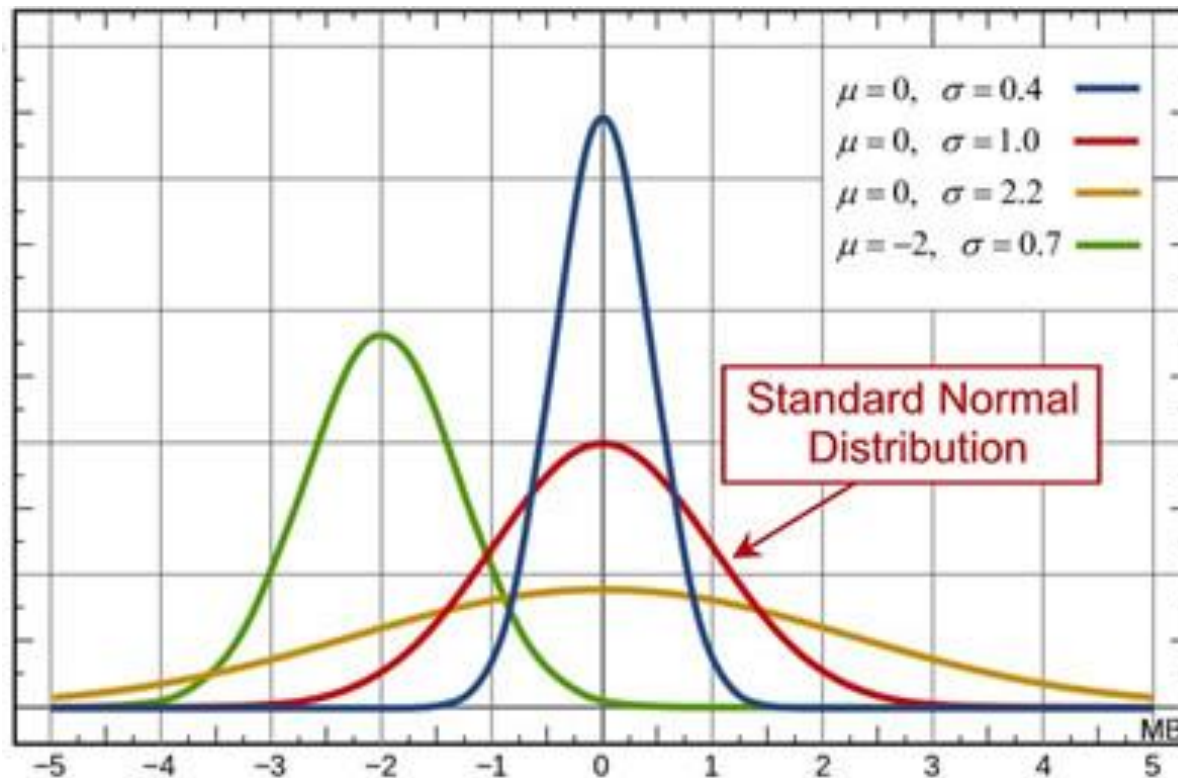
### توزیع نرمال □

تابع چگالی:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\mu$  = mean of  $x$

$\sigma$  = standard deviation of  $x$



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

گروه دایچه | dayche.com

# تئوری احتمال

## توزیع های احتمالی

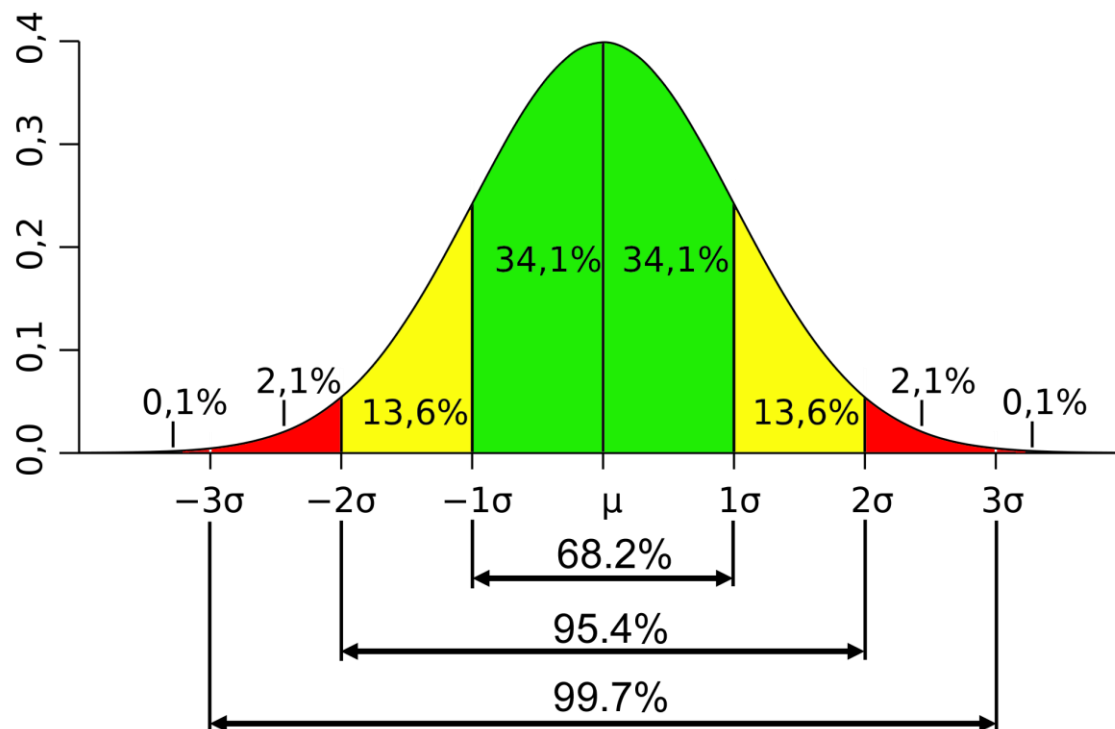
### توزیع نرمال □

تابع چگالی:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\mu$  = mean of  $x$

$\sigma$  = standard deviation of  $x$



تولید محتوا: زهرا ذوالقدر

daychegroup 

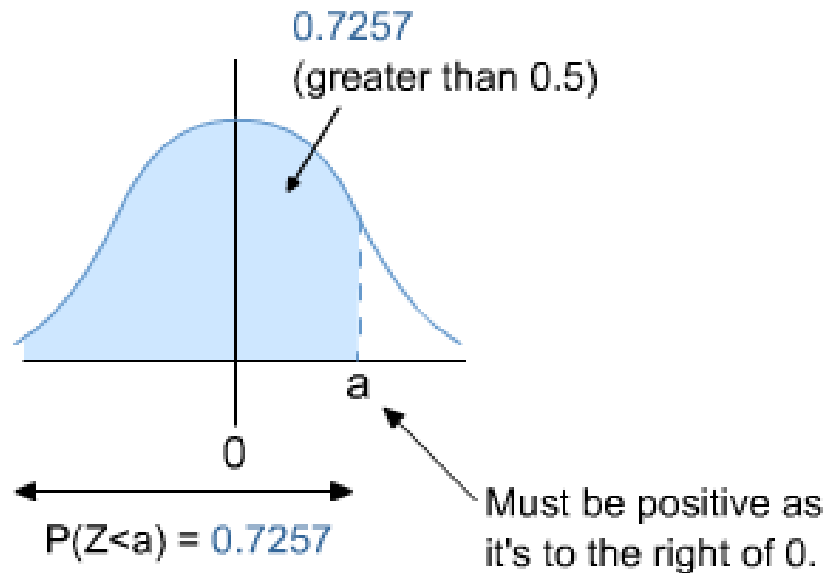
daychegroup 

گروه دایچه | dayche.com 



### توزیع نرمال □

محاسبه چندک و احتمال تجمعی



Therefore:

$$P(Z < a) = 0.7257$$

Means:

$$\Phi(a) = 0.7257$$

So:

$$a = 0.6$$

تولید محتوا: زهرا ذوالقدر

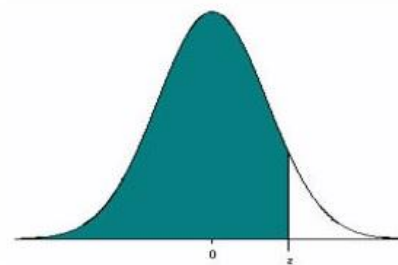
daychegroup

daychegroup

dayche.com | گروه دایکه



Table of Standard Normal Probabilities for Positive Z-scores



توزیع نرمال □

نحوه استفاده از جدول توزیع نرمال استاندارد

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916

تولید محتوا: زهرا ذوالقدر

daychegroup

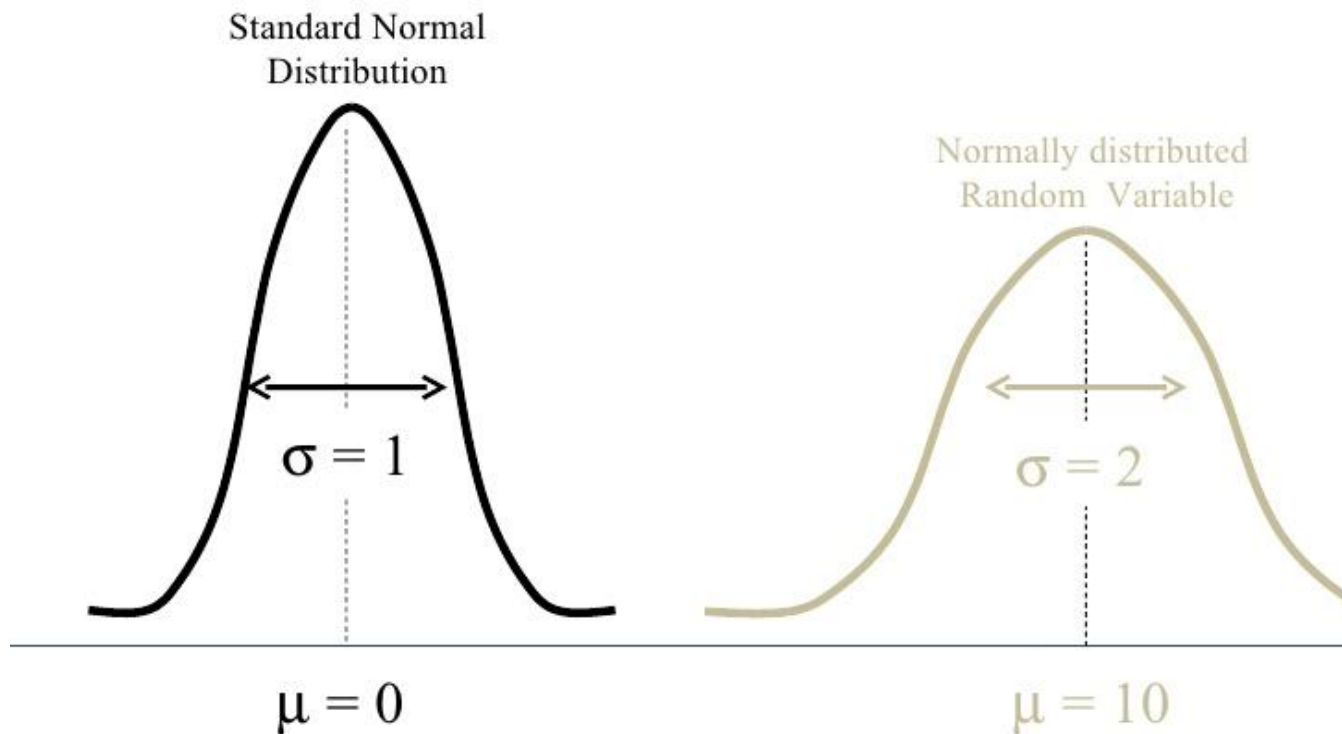
daychegroup

dayche.com | گروه دایکه

We use the formula for Z transformation:  $Z = \frac{X - \mu}{\sigma}$

توزیع نرمال □


تبدیل خطی توزیع را حفظ می کند.



تولید محتوا: زهرا ذوالقدر

daychegroup 

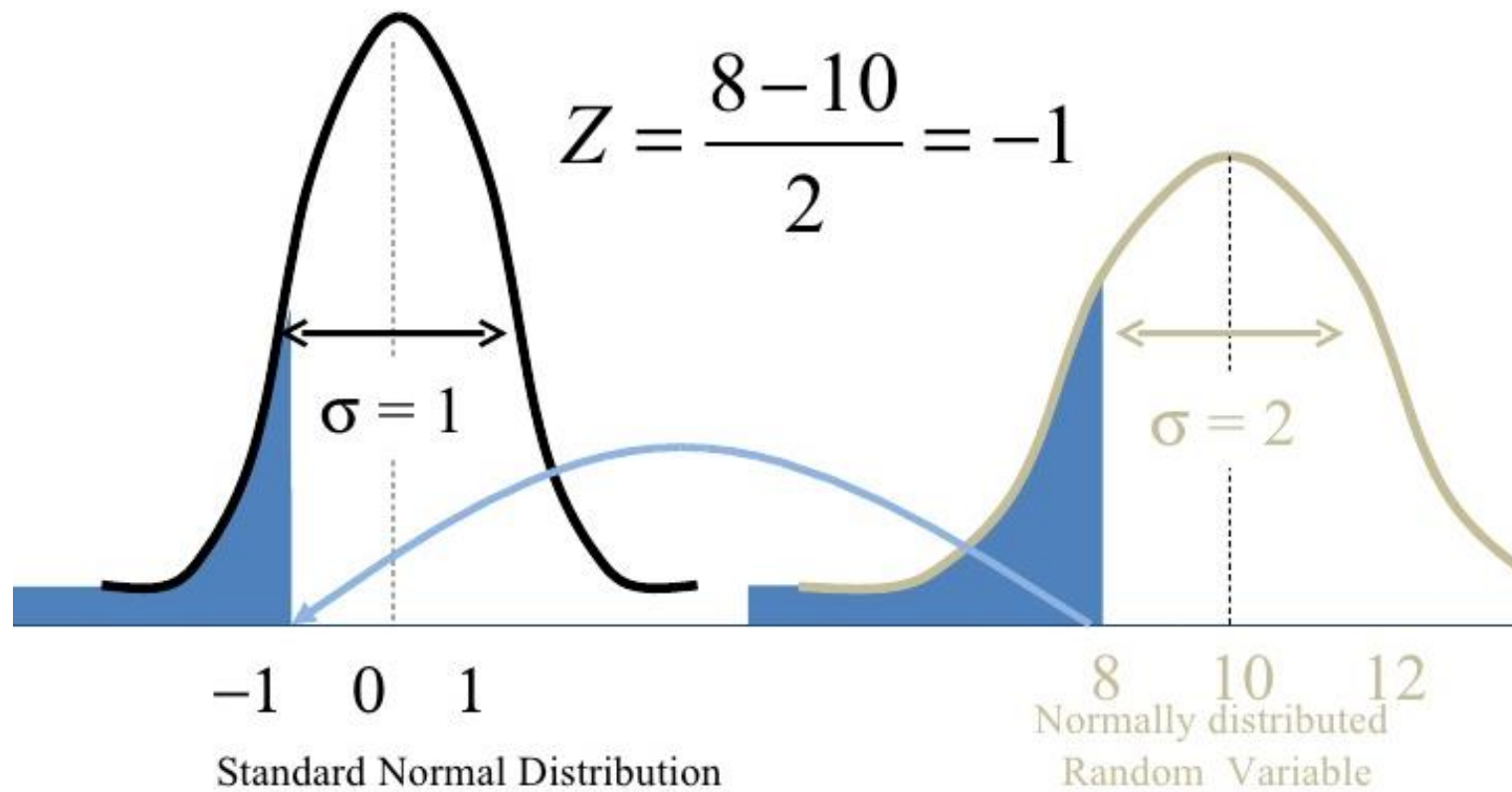
daychegroup 

dayche.com | گروه دایچه 

# نئوری احتمال

## توزیع های احتمالی


توزیع نرمال □



تولید محتوا: زهرا ذوالقدر

daychegroup 

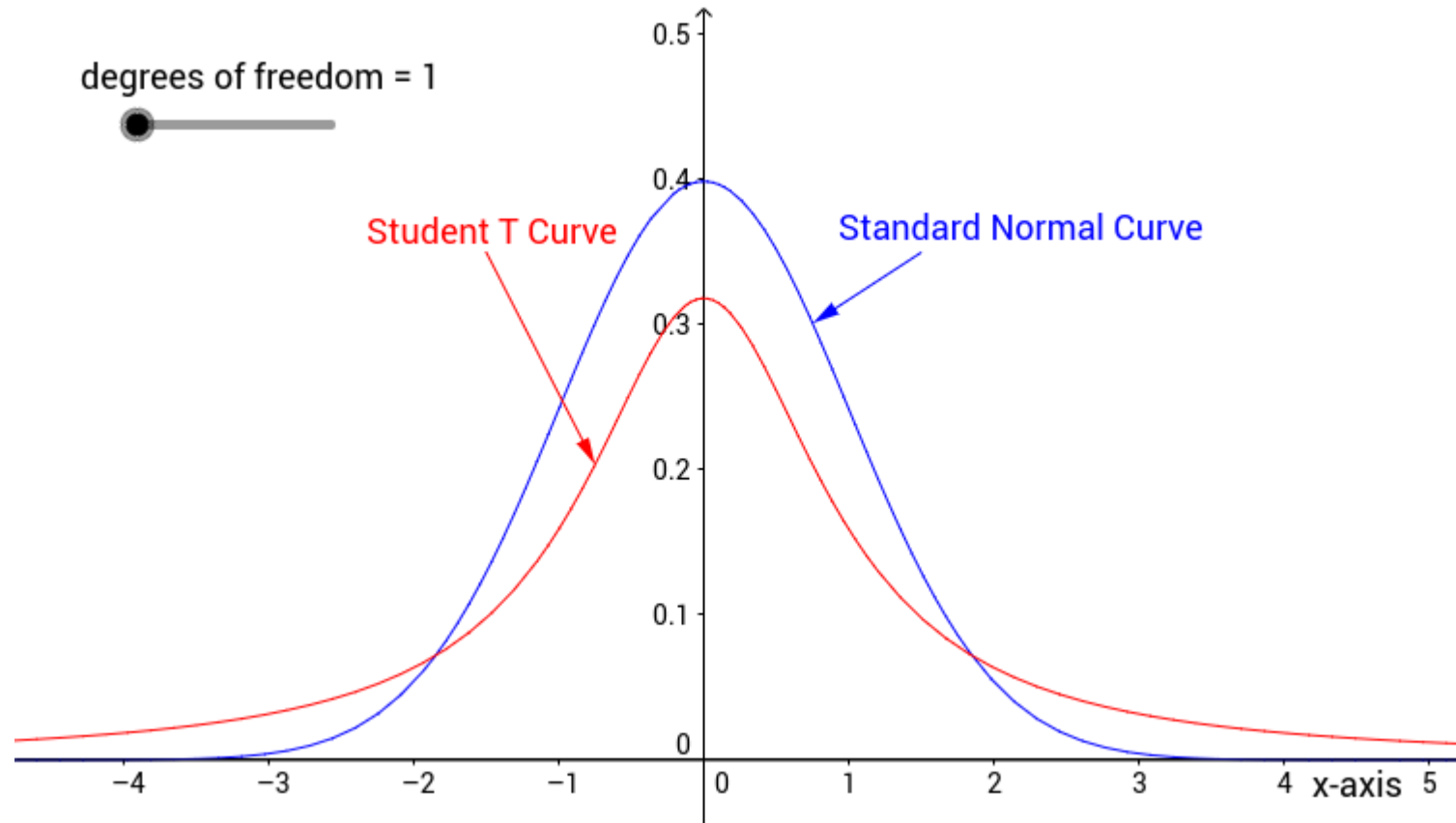
daychegroup 

dayche.com | گروه دایچه 

# تئوری احتمال

## توزیع های احتمالی


توزیع t



تولید محتوا: زهرا ذوالقدر

daychegroup 

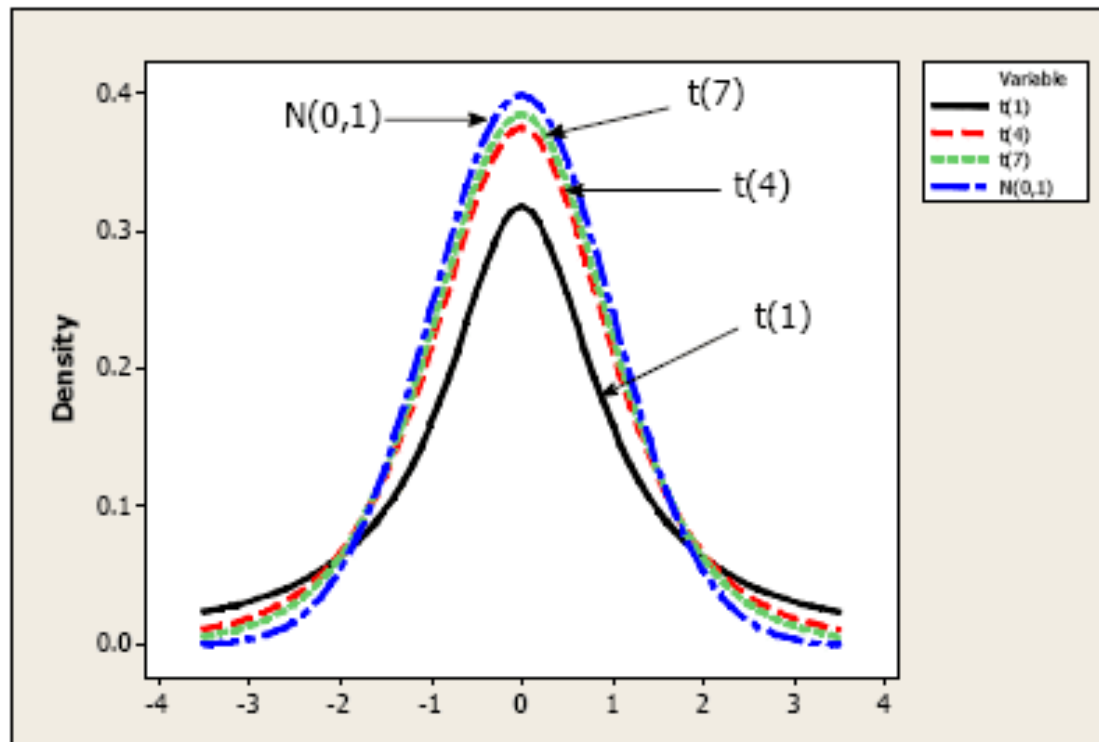
daychegroup 

dayche.com | گروه دایچه 

### توزیع t □

تابع چگالی:


$$f(x) = \frac{1}{\sqrt{\pi v}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{x^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}$$



تولید محتوا: زهرا ذوالقدر

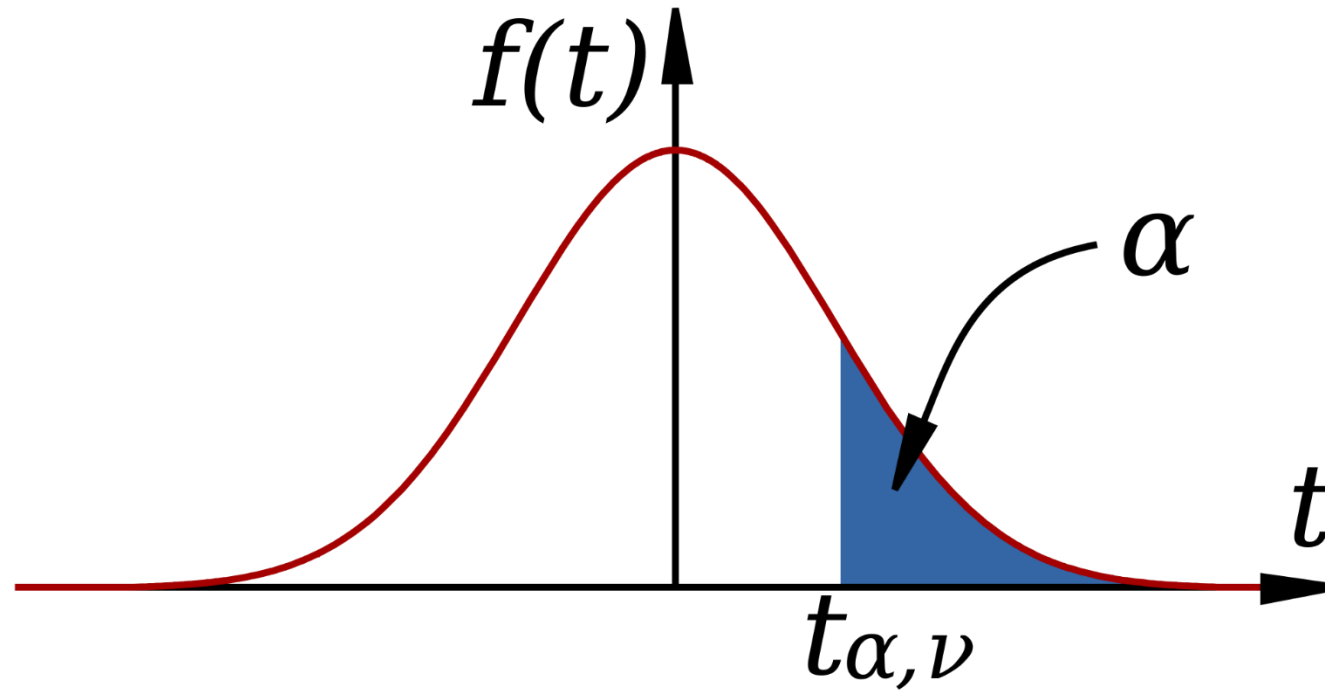
daychegroup 

daychegroup 

گروه دایچه | dayche.com 




توزیع  $t$  □



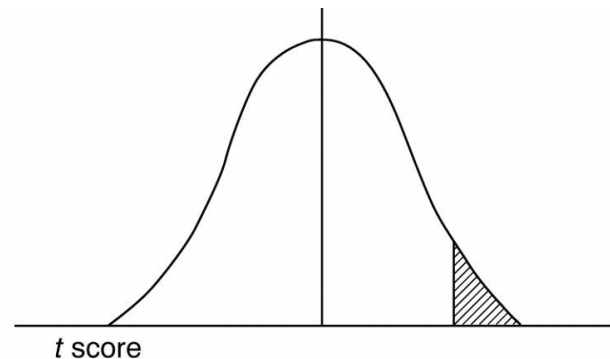
تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

### توزیع t




نحوه استفاده از جدول توزیع t

df \ p	0.1	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.683	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.160	2.650	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
24	1.318	1.711	2.064	2.492	2.797

تولید محتوا: زهرا ذوالقدر

daychegroup 

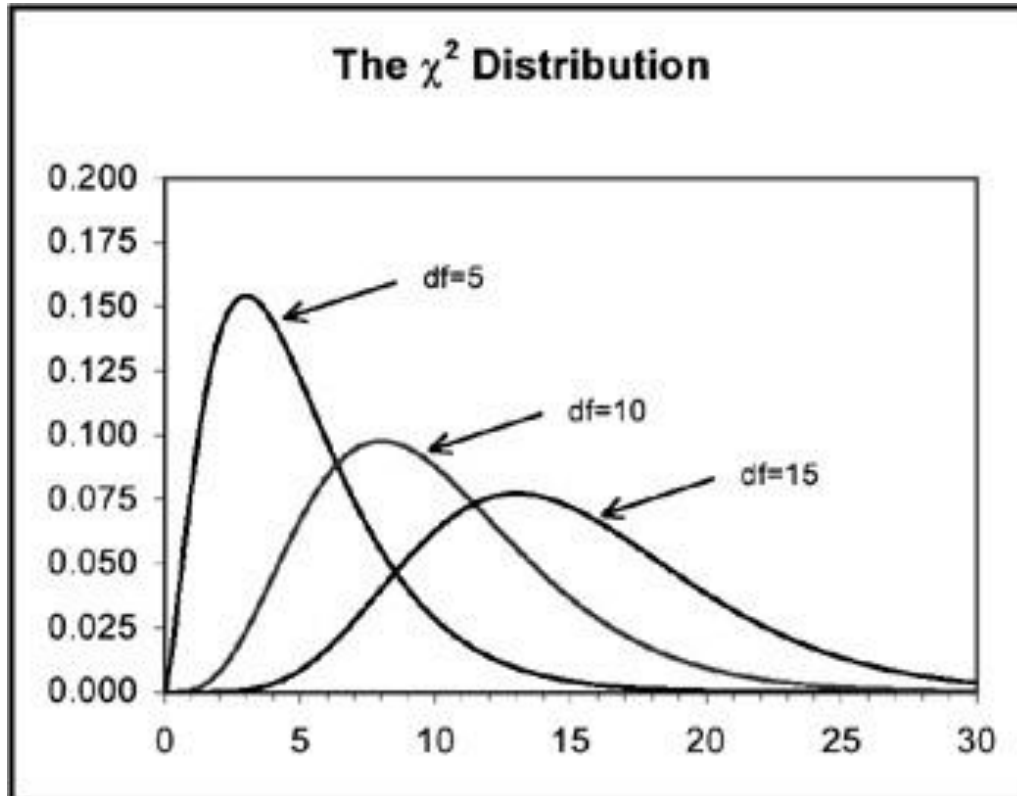
daychegroup 

dayche.com | گروه دایچه 

### توزیع کای-دو □

تابع چگالی:


$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}$$



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

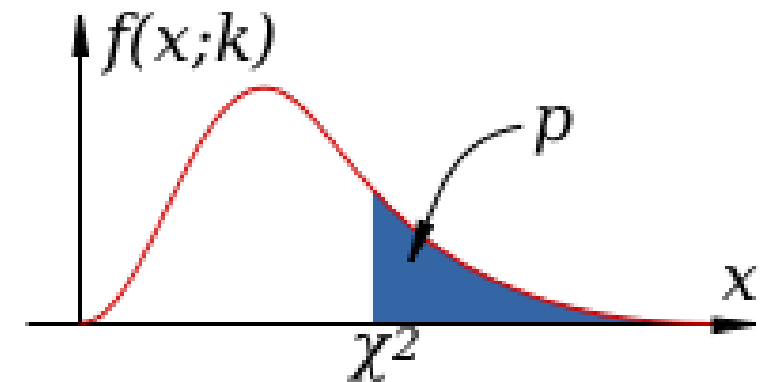
dayche.com | گروه دایکه 

## Chi-Square table

39

Degrees of Freedom	Level of Significance								
	0.99	0.95	0.9	0.75	0.2	0.1	0.05	0.01	0.001
1	0.000157	0.004	0.016	0.102	1.642	2.706	3.841	6.635	10.827
2	0.0201	0.103	0.211	0.575	3.219	4.605	5.991	9.210	13.815
3	0.115	0.352	0.584	1.213	4.642	6.251	7.815	11.345	16.266
4	0.297	0.711	1.064	1.923	5.989	7.779	9.488	13.277	18.466
5	0.554	1.145	1.610	2.675	7.289	9.236	11.070	15.086	20.515
6	0.872	1.635	2.204	3.455	8.558	10.645	12.592	16.812	22.457
7	1.239	2.167	2.833	4.255	9.803	12.017	14.067	18.475	24.321
8	1.647	2.733	3.490	5.071	11.030	13.362	15.507	20.090	26.124
9	2.088	3.325	4.168	5.899	12.242	14.684	16.919	21.666	27.877
10	2.558	3.940	4.865	6.737	13.442	15.987	18.307	23.209	29.588
11	3.053	4.575	5.578	7.584	14.631	17.275	19.675	24.725	31.264
12	3.571	5.226	6.304	8.438	15.812	18.549	21.026	26.217	32.909
13	4.107	5.892	7.041	9.299	16.985	19.812	22.362	27.688	34.527
14	4.600	6.571	7.790	10.165	18.151	21.064	23.685	29.141	36.124
15	5.229	7.261	8.547	11.037	19.311	22.307	24.996	30.578	37.698
16	5.812	7.962	9.312	11.912	20.465	23.542	26.296	32.000	39.252
17	6.408	8.672	10.085	12.792	21.615	24.769	27.587	33.409	40.791
18	7.015	9.390	10.865	13.675	22.760	25.989	28.869	34.805	42.312

نحوه استفاده از جدول توزیع کای-دو



تولید محتوا: زهرا ذوالقدر

daychegroup

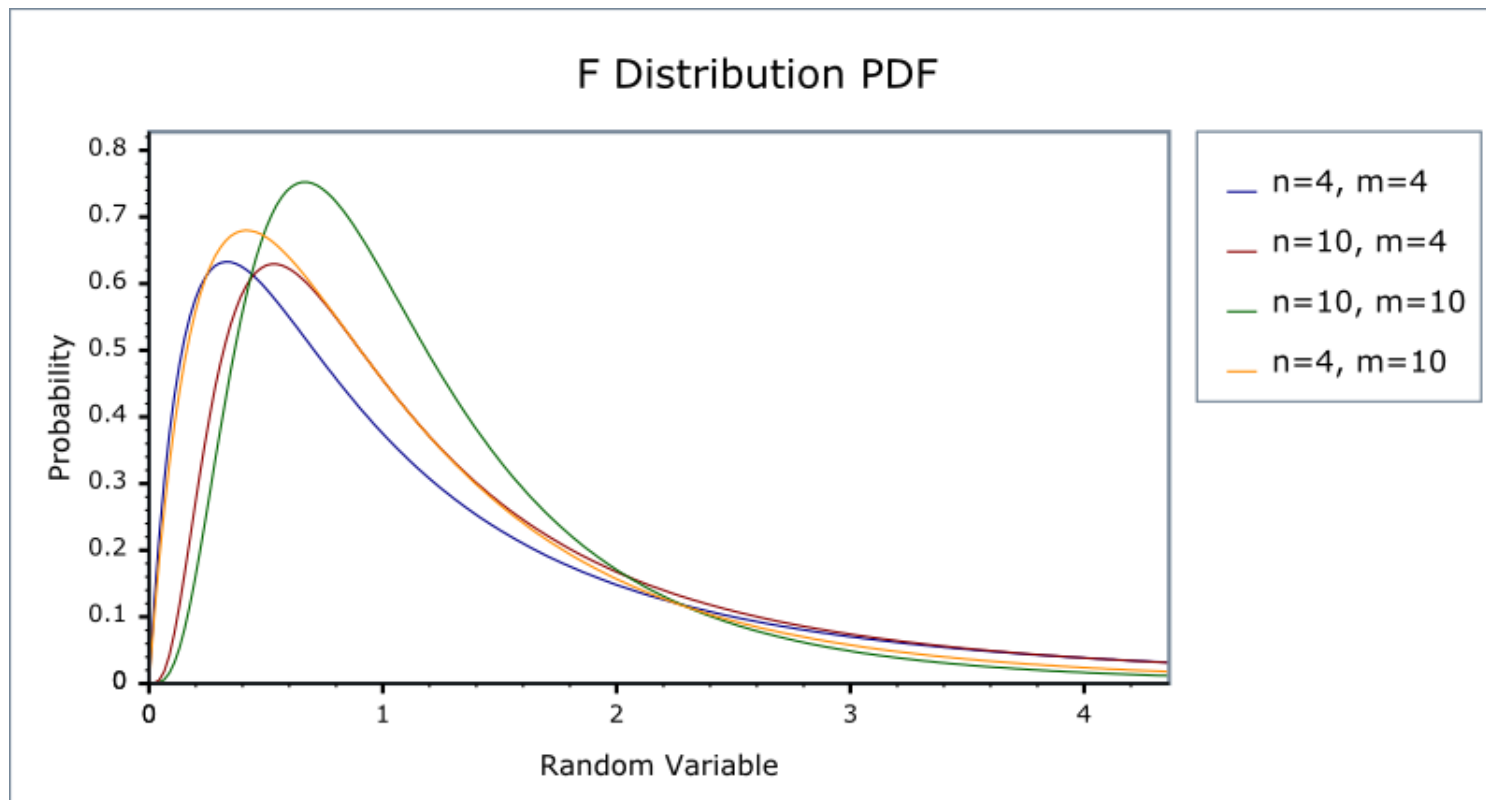
daychegroup

dayche.com | گروه دایچه

### F توزیع □

تابع چگالی:


$$f(x) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right) \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} x^{\left(\frac{\nu_1}{2} - 1\right)}}{\Gamma\left(\frac{\nu_1}{2}\right) \Gamma\left(\frac{\nu_2}{2}\right) (\nu_2 + \nu_1 x)^{\left(\frac{\nu_1 + \nu_2}{2}\right)}}$$



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تئوری احتمال

## توزیع های احتمالی

F توزیع □

F Table for  $\alpha = .05$




نحوه استفاده از جدول توزیع F

	$df_1=1$	2	3	4	5	6	7	8	9	10
$df_2=1$	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

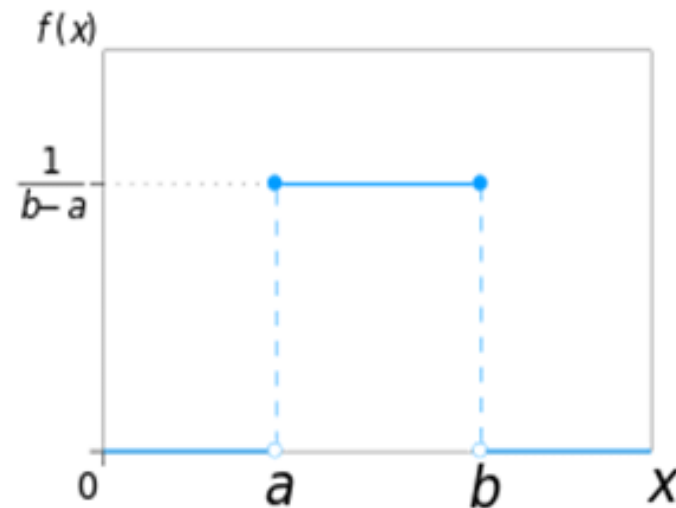
# تئوری احتمال

## توزیع های احتمالی

### توزیع یکنواخت

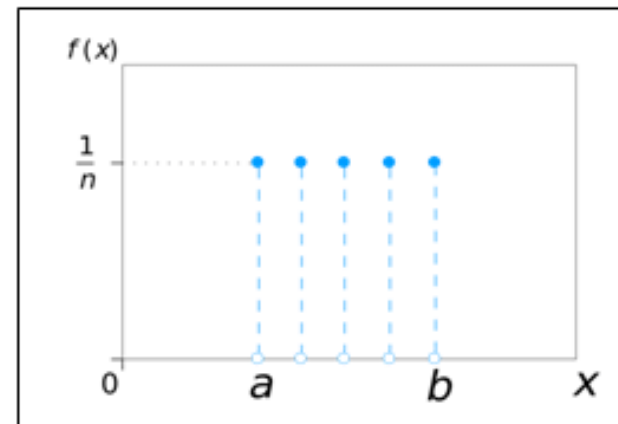
توزیعی که چگالی احتمال در همه نقاط یکسان است.

$$f(x) = \frac{1}{b-a} \quad a < x < b$$



(a) Continuous Uniform Distribution

$$f(x) = \frac{1}{n} \quad \text{for } n \text{ events}$$




(b) Discrete Uniform Distribution

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

گروه دایچه | dayche.com 

# تئوری احتمال

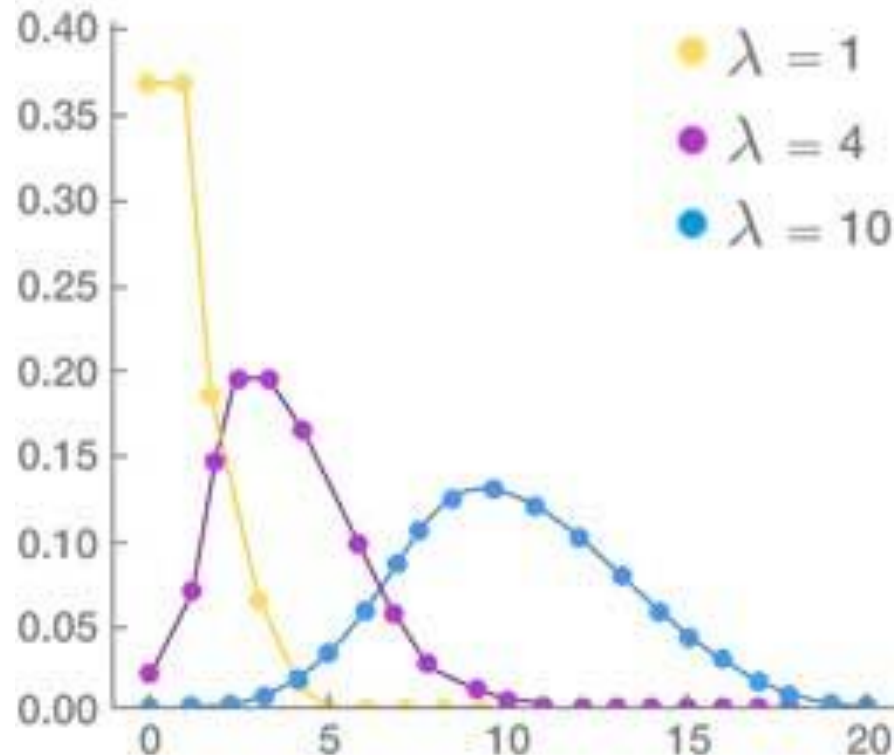
## توزیع های احتمالی

### توزیع پواسن □

تابع چگالی:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$E(x) = \lambda$$




○ شمارش رخداد اتفاقات نادر در زمان و مکان مشخص

○ ارتباط با توزیع دو جمله ای

تولید محتوا: زهرا ذوالقدر

daychegroup 

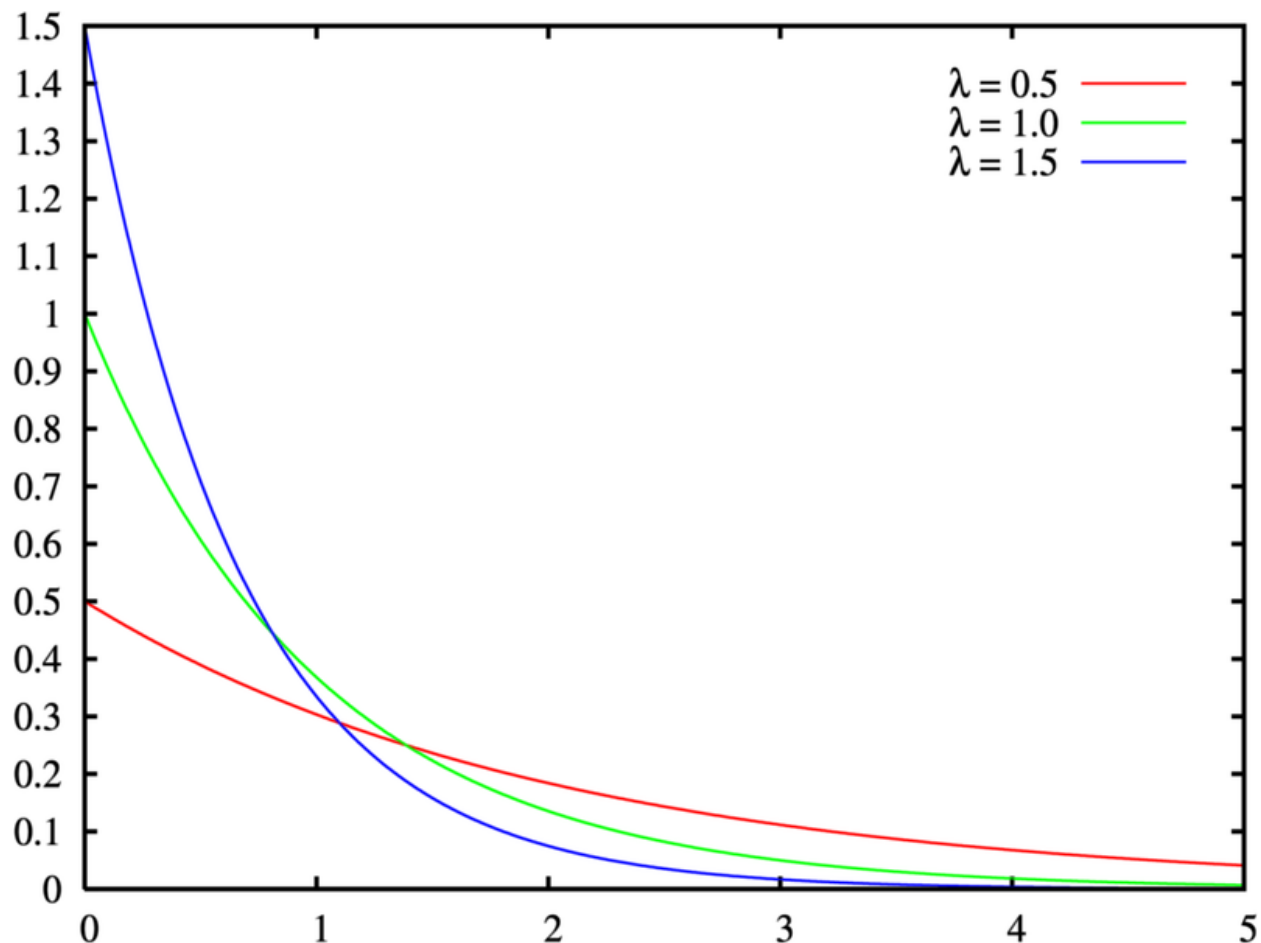
daychegroup 

dayche.com | گروه دایچه 



# تئوری احتمال

## توزیع های احتمالی



□ توزیع نمایی

تابع چگالی:

$$f(x) = \lambda x e^{-\lambda x}$$

$$E(x) = \frac{1}{\lambda}$$

○ زمان تا رخداد اتفاق خاص

○ خاصیت عدم حافظه

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تئوری احتمال

## توزیع های احتمالی

### ارتباط توزیع نمایی و پواسن □

توزیع نمایی ○

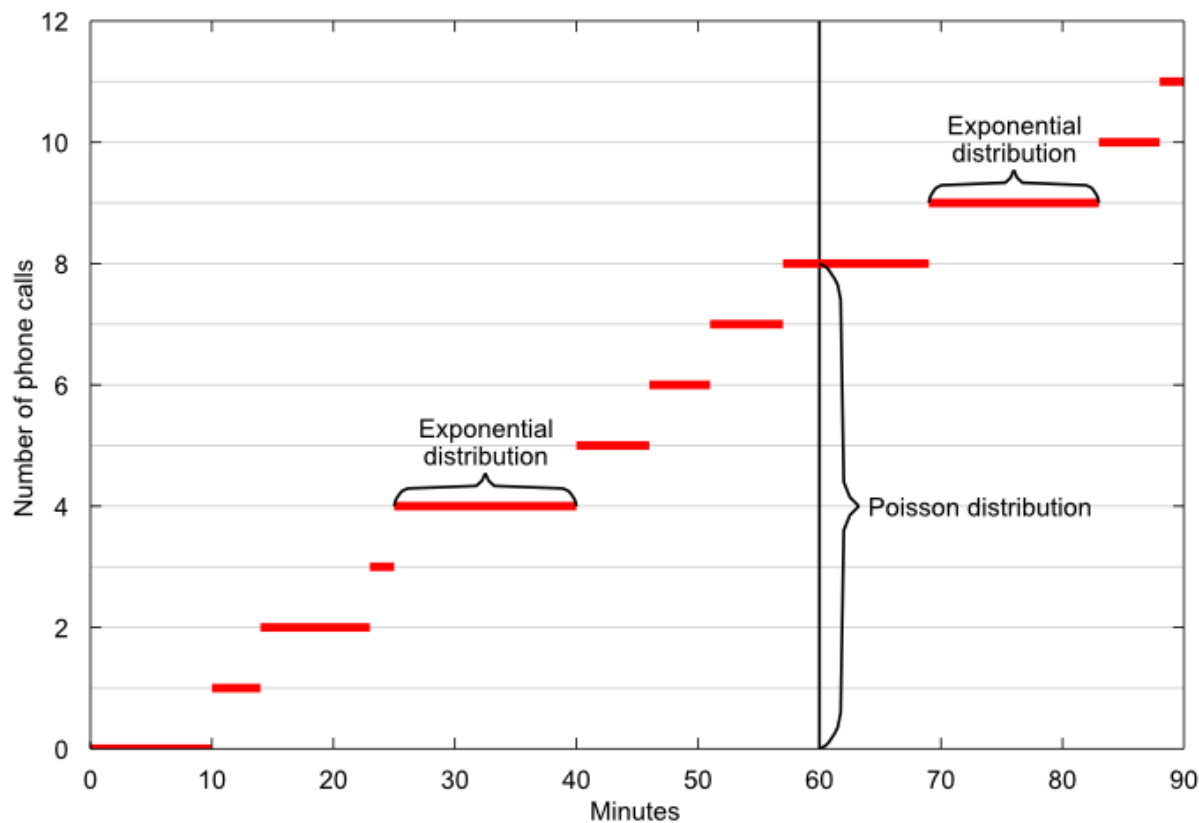
$$f(x) = \lambda e^{-\lambda x}$$

$$E(x) = \frac{1}{\lambda}$$

توزیع پواسن ○

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$E(x) = \lambda$$



تولید محتوا: زهرا ذوالقدر

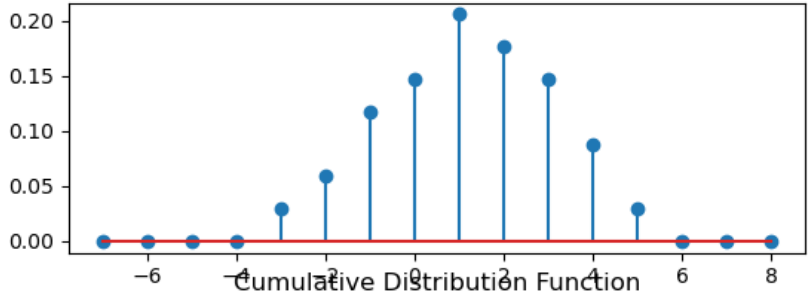
daychegroup 

daychegroup 

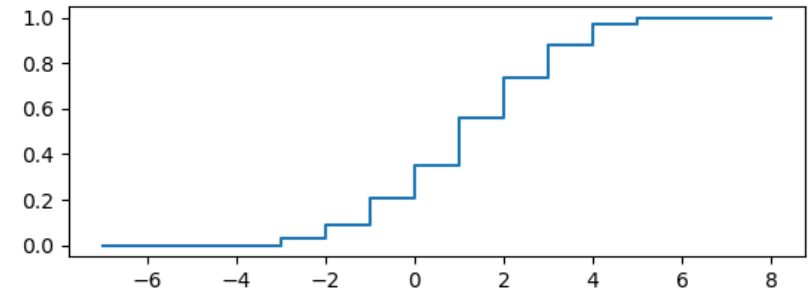
گروه دایچه | dayche.com 

تابع توزیع تجمعی (Cumulative Distribution Function) □

Probability Mass Function

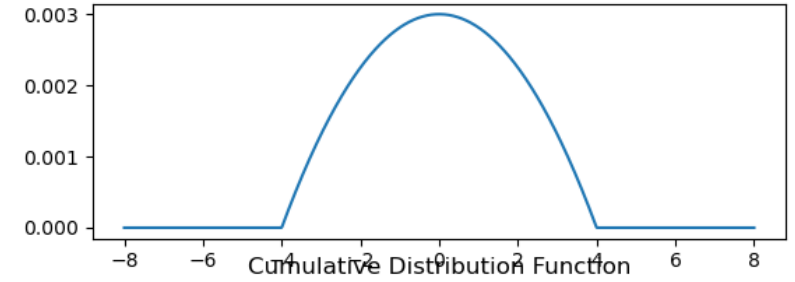


Cumulative Distribution Function

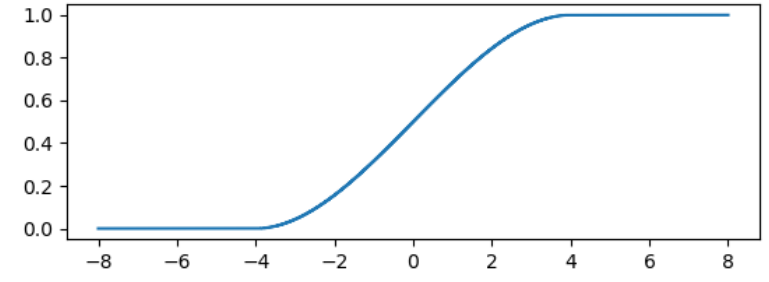


$$F(x) = P(X \leq x)$$

Probability Density Function



Cumulative Distribution Function



$$F_X(t) = \begin{cases} \sum_{\substack{x \in A \\ x \leq t}} p(x), & \text{r. v. } X \text{ is discrete} \\ \int_{-\infty}^t f(x) dx, & \text{r. v. } X \text{ is continuous} \end{cases}$$

تولید محتوا: زهرا ذوالقدر

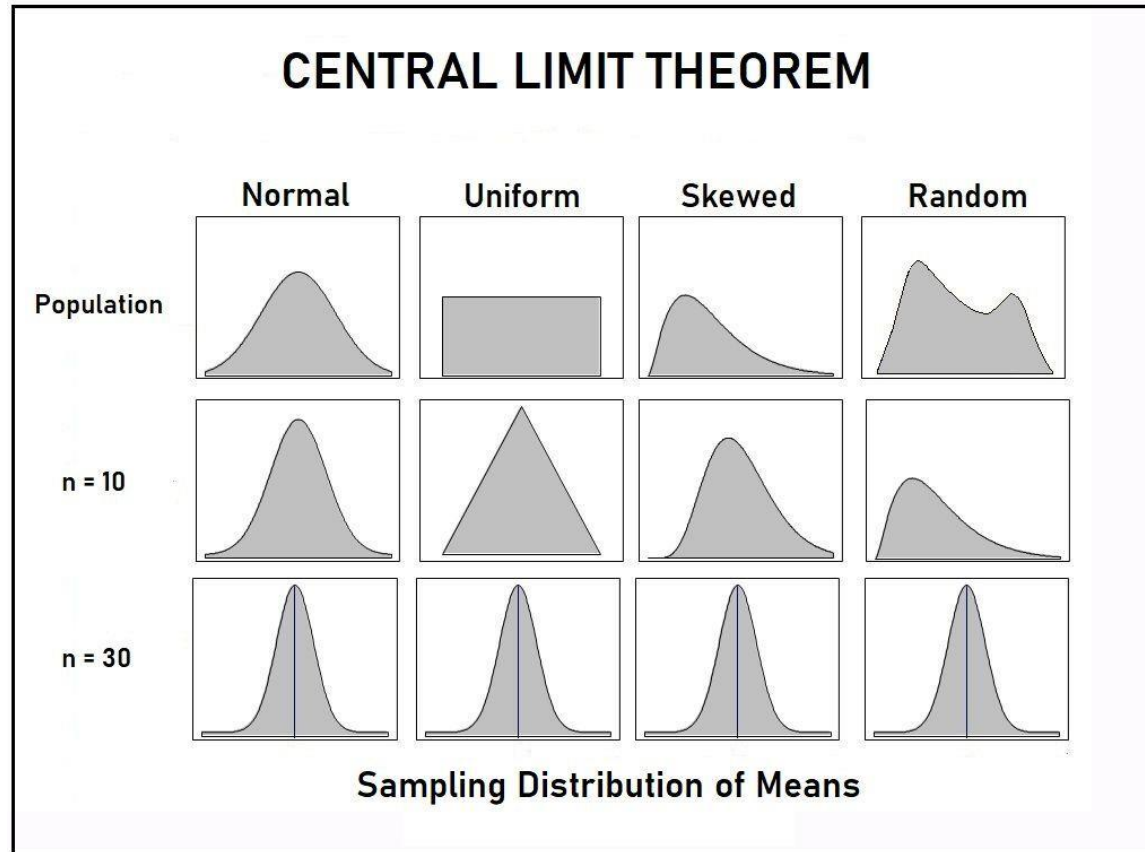
daychegroup

daychegroup

گروه دایکه | dayche.com

### □ قضیه حد مرکزی (The Central Limit Theorem)


متغیر تصادفی  $X$  هر توزیعی داشته باشد، در صورتی که نمونه های به اندازه کافی بزرگ از این توزیع استخراج کنیم توزیع تقریبی  $\bar{X}$  نرمال خواهد بود.



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

# تئوری احتمال

## توزیع های احتمالی

### امید ریاضی (Expected Value) □

- یک قمارباز ممکن است علاقه مند به دانستن متوسط برد خود در هر بازی باشد.
- صاحب یک کسب و کار مایل است متوسط سود روی محصول یا خدمات خود را بداند.
- خریداران باید بین کالای حراجی یا کالای تضمینی انتخاب کنند.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

### □ امید ریاضی (Expected Value)

○ یک قمارباز ممکن است علاقه مند به دانستن متوسط برد خود در هر بازی باشد.

**سوال:** بازی زیر به شما پیشنهاد می شود:

دو تاس را پرتاب کنید. اگر مجموع خال ها 6، 7 یا 8 بود پیشنهاد دهنده برنده خواهد بود و باید 200 هزار تومان پرداخت کنید. اگر مجموع خال ها 2، 3، 4، 5، 9، 10، 11 یا 12 بود شما برنده خواهید بود و 100 هزار تومان دریافت خواهید کرد. آیا بازی می کنید؟

# تئوری احتمال

## توزیع های احتمالی

□ امید ریاضی (Expected Value)

حل:

آزمایش تصادفی: پرتاب دو تاس

فضای نمونه:  $S = \{(1,1) (1,2) \dots (1,6) \dots (6,6)\}$


برآمد آزمایش	P	X
مجموع خال ها 6، 7 یا 8	$16/36$	-200
مجموع خال ها 2، 3، 4، 5، 9، 10، 11 یا 12	$20/36$	100

$$16/36(-200) + 20/36(100) = -100/3$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

### □ امید ریاضی (Expected Value)

○ صاحب یک کسب و کار مایل است متوسط سود روی محصول یا خدمات خود را بداند.

**سوال:** یک مرد 52 ساله خود را بیمه عمر می کند. بر اساس تجربه بیمه گر احتمال اینکه مرد 52 ساله برای یک سال آینده زنده بماند 0.995 است. شرکت بیمه مبلغ 208 هزار تومان از بیمه شونده در مقابل بیمه 40 میلیون تومانی بیمه عمر دریافت می کند. اگر  $X$  نمایانگر سود شرکت باشد متوسط سود شرکت بیمه از این فرد چقدر است؟



# تئوری احتمال

## توزیع های احتمالی

امید ریاضی (Expected Value) □

حل:

برآمد آزمایش	P	X
مرد بیمه شده تا پایان سال بیمه زنده می ماند	0.995	208000
مرد بیمه شده تا پایان سال بیمه زنده نمی ماند	0.005	208000-40000000

$$208000 (0.995) + (-39792000)(0.005) = 8000$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

# تئوری احتمال

## توزیع های احتمالی

### امید ریاضی (Expected Value) □


○ خریداران باید بین کالای حراجی یا کالای تضمینی انتخاب کنند.

**سوال:** قیمت یک کالا در حراجی 30 هزار تومان است و با احتمال 0.3 معیوب. قیمت تضمینی کالا 40 هزار تومان است. کدامیک را ترجیح می دهید؟

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

# تئوری احتمال

## توزیع های احتمالی

امید ریاضی (Expected Value) □

حل:

X : مقدار سود


برآمد آزمایش	P	X
کالای خریداری شده سالم است	0.7	10000
کالای خریداری شده معیوب است	0.3	-30000

$$E(X) = 10000(0.7) + (-30000)(0.3) = -2000$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

### امید ریاضی (Expected Value) □

- در واقع امید ریاضی میانگین وزنی  $X$  با وزن های احتمال است.
- امید ریاضی در سرمایه گذاری، اقتصاد و امور مالی و همچنین تحلیل سناریو کاربردهای زیادی دارد.
- امید ریاضی میانگین جامعه است.

$$E[X] = \sum_i x_i f(x_i)$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

# Statistical Analysis

تحليل آماری

گروه دایچه . dayche.com





## □ آمار توصیفی Descriptive Statistics

خلاصه سازی داده های نمونه با روش هایی نظیر جداول فراوانی، نمودارها و شاخص های آماری.


## □ آمار استنباطی Inferential Statistics

مجموعه ای از روش های آماری که بر اساس اطلاعات حاصل از نمونه استنباط هایی در باره جامعه آماری به دست می آید.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



## مقیاس های اندازه گیری

- مقیاس های اسمی (Nominal Scales). مانند رنگ محصول و گروه خونی
- مقیاس های رتبه ای (Ordinal Scales). مانند سطح تحصیلات و سطح رضایت
- مقیاس های عددی (Numerical Scales): سن و فشار خون

# تحلیل آماری

## آمار توصیفی برای متغیرهای کیفی

### شخص های جدول فراوانی


- فراوانی
- فراوانی تجمعی
- فراوانی نسبی
- فراوانی نسبی تجمعی

شخص مد (نما): رده ای که بیشترین فراوانی را دارد.

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



# تحلیل آماری

## آمار توصیفی برای متغیرهای کیفی

### شاخص های جدول فراوانی

- فراوانی
- فراوانی تجمعی
- فراوانی نسبی
- فراوانی نسبی تجمعی


<u>No. of Siblings</u> <b><i>X</i></b>	<u>Frequency</u> <b><i>f</i></b>	<u>Cumulative Frequency</u> <b><i>cf</i></b>	<u>Relative Frequency</u> <b><i>rf</i></b>	<u>Cumulative Relative Frequency</u> <b><i>crf</i></b>
0	3	3	0.11	0.11
1	7	10	0.25	0.36
2	10	20	0.36	0.71
3	3	23	0.11	0.82
4	4	27	0.14	0.96
5	0	27	0.00	0.96
6	1	28	0.04	1.00
Total	28			

مثال: مقدار مد در جدول مقابل: 2

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تحلیل آماری

## آمار توصیفی برای متغیرهای کمی

### □ شاخص های مرکزی

میانگین: حاصل جمع اعداد تقسیم بر تعداد

میانه: نقطه وسط داده ها

نما: داده با بیشترین تکرار

مثال: 1 1 2 3 3 4 5 6 8 38

میانگین: 7


میانه: 3

مد: 3

تولید محتوا: زهرا ذوالقدر

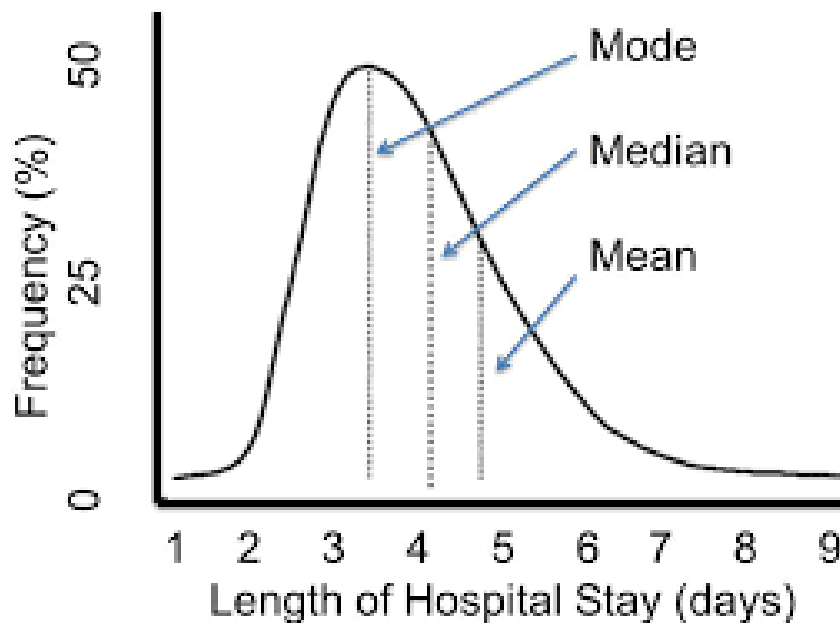
daychegroup 

daychegroup 

dayche.com | گروه دایکه 

# تحلیل آماری

آمار توصیفی برای متغیرهای کمی



شاخص های مرکزی □

1 1 2 3 3 4 5 6 38



Median



Mean

تولید محتوا: زهرا ذوالقدر

daychegroup

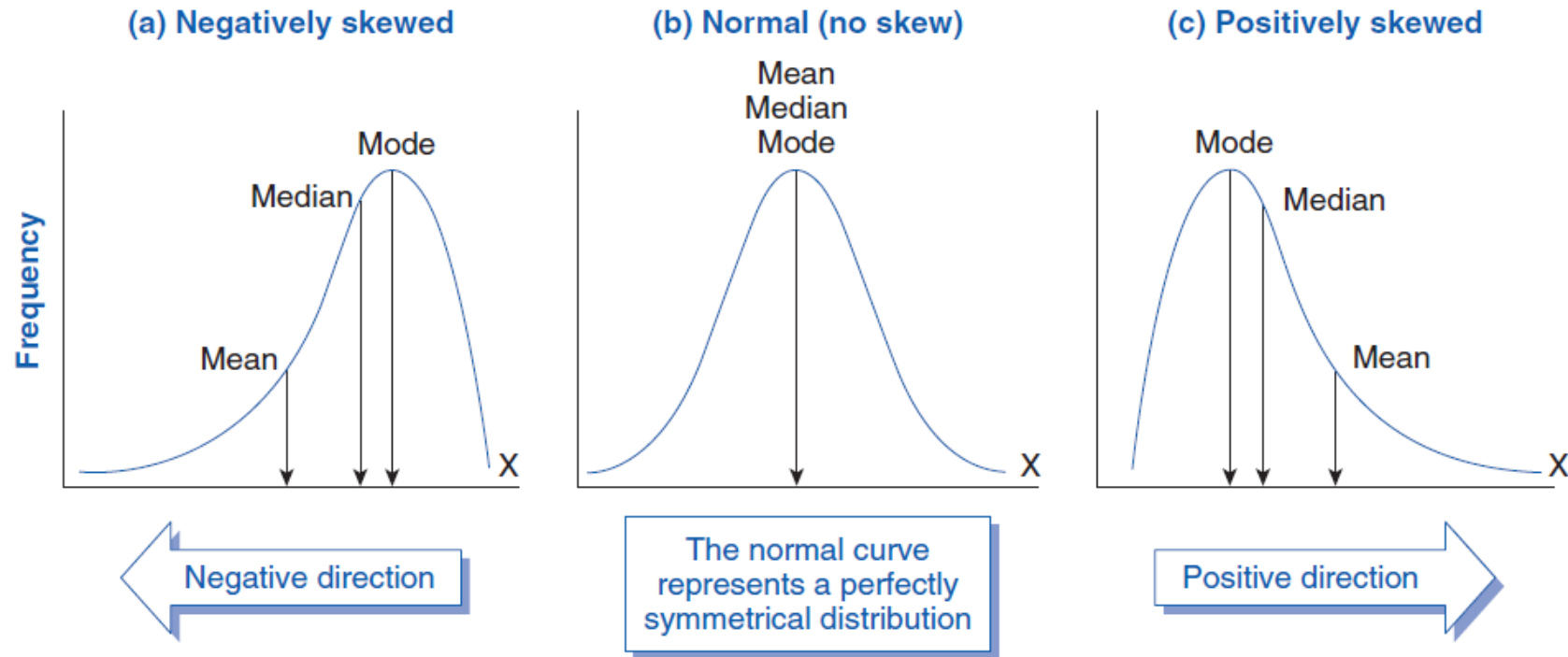
daychegroup

dayche.com | گروه دایچه

# تحلیل آماری

## آمار توصیفی برای متغیرهای کمی

چولگی □



تولید محتوا: زهرا ذوالقدر

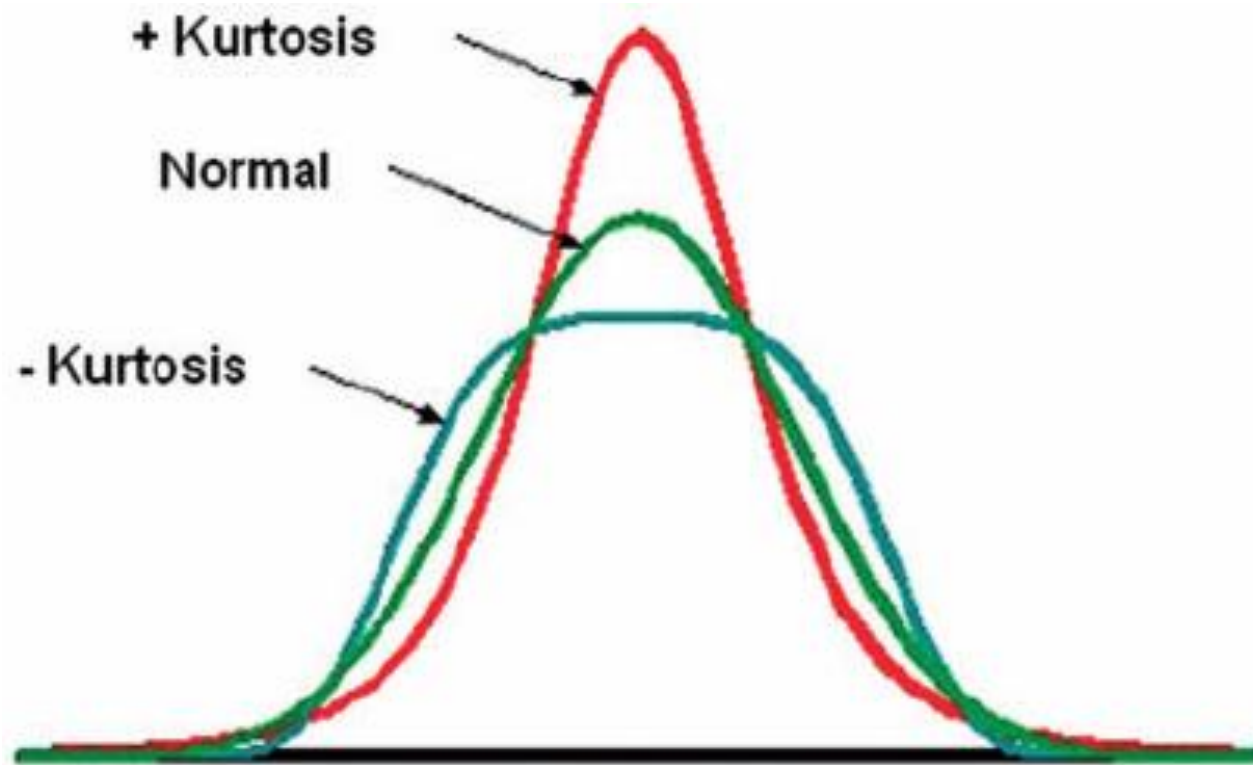
daychegroup

daychegroup

dayche.com | گروه دایچه

# تحلیل آماری

آمار توصیفی برای متغیرهای کمی




کشیدگی □

تولید محتوا: زهرا ذوالقدر

daychegroup 

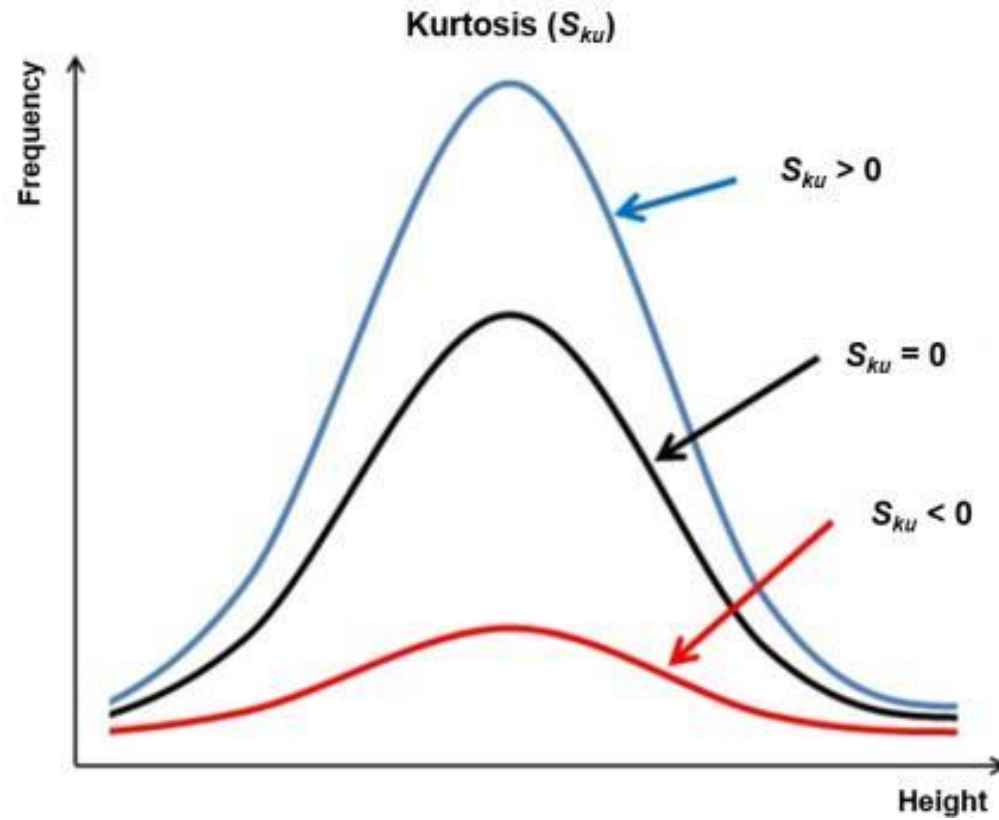
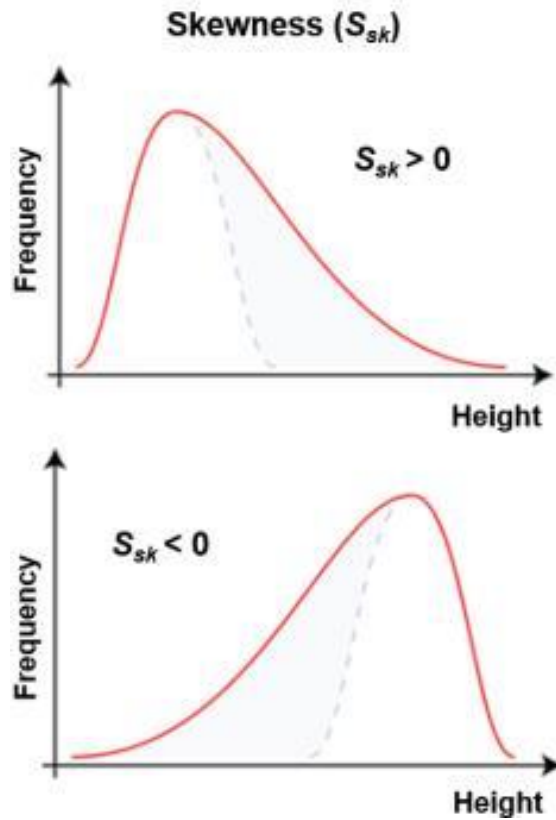
daychegroup 

dayche.com | گروه دایکه 

# تحلیل آماری

آمار توصیفی برای متغیرهای کمی


چولگی و کشیدگی □



تولید محتوا: زهرا ذوالقدر

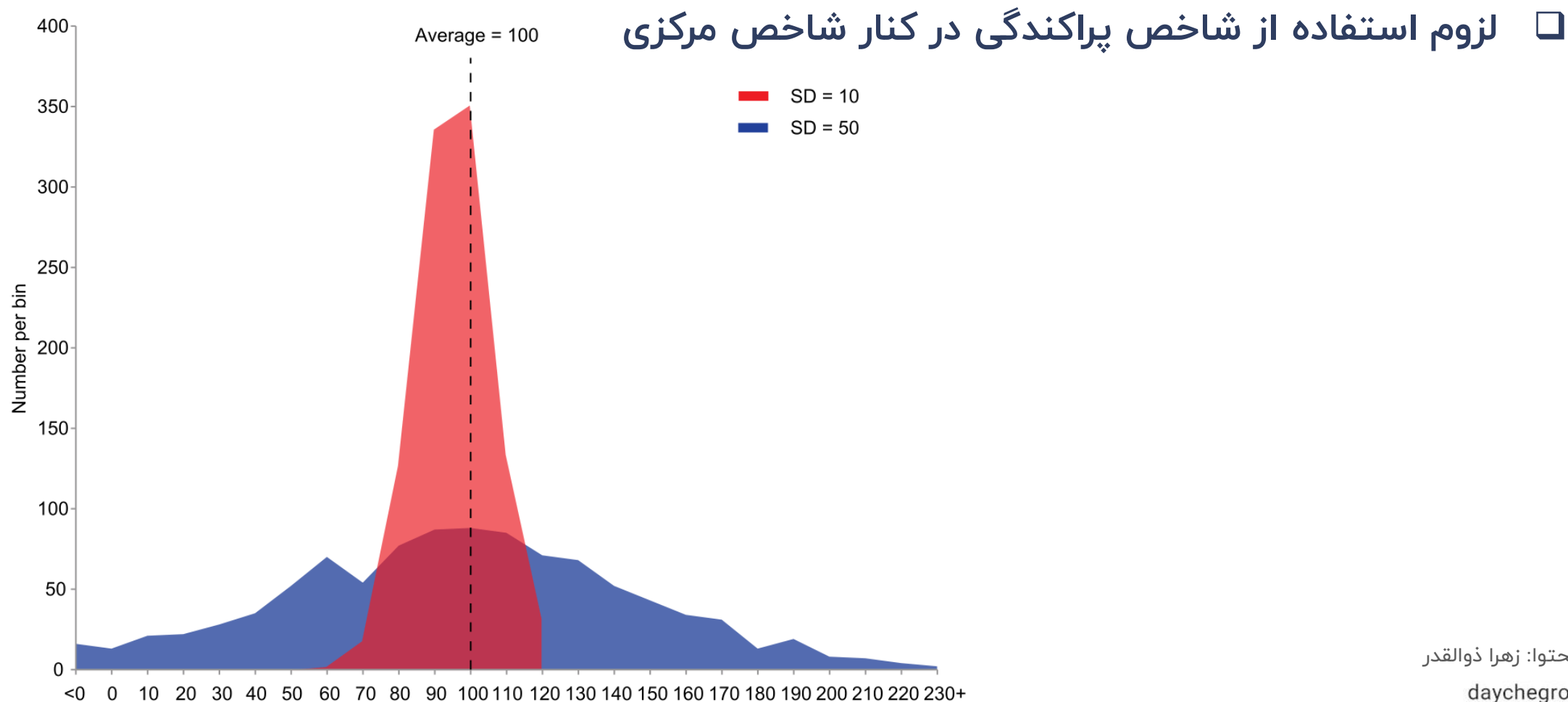
daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تحلیل آماری

آمار توصیفی برای متغیرهای کمی



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

گروه دایکه | dayche.com

# تحلیل آماری

## آمار توصیفی برای متغیرهای کمی

### □ دیگر شاخص های پراکندگی

دامنه داده ها: اختلاف کوچکترین و بزرگترین داده

واریانس: میانگین انحراف داده ها از میانگین


انحراف معیار: جذر واریانس

دامنه میان چارکی: اختلاف چارک اول و سوم داده ها

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 



# تحلیل آماری

آمار توصیفی برای متغیرهای کمی

❑ دیگر شاخص های پراکندگی

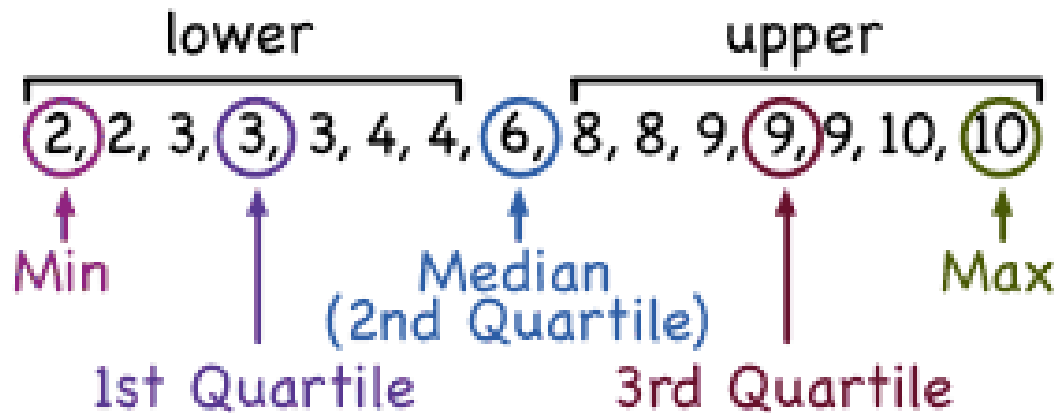
دامنه میان چارکی: اختلاف چارک اول و سوم داده ها

Find Min, Max, Median & Quartiles

**Glasses of Water Drank Per Day:**

2, 4, 3, 9, 8, 3, 10, 6, 2, 9, 3, 9, 8, 4, 10

✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓



تولید محتوا: زهرا ذوالقدر

daychegroup

daychegroup

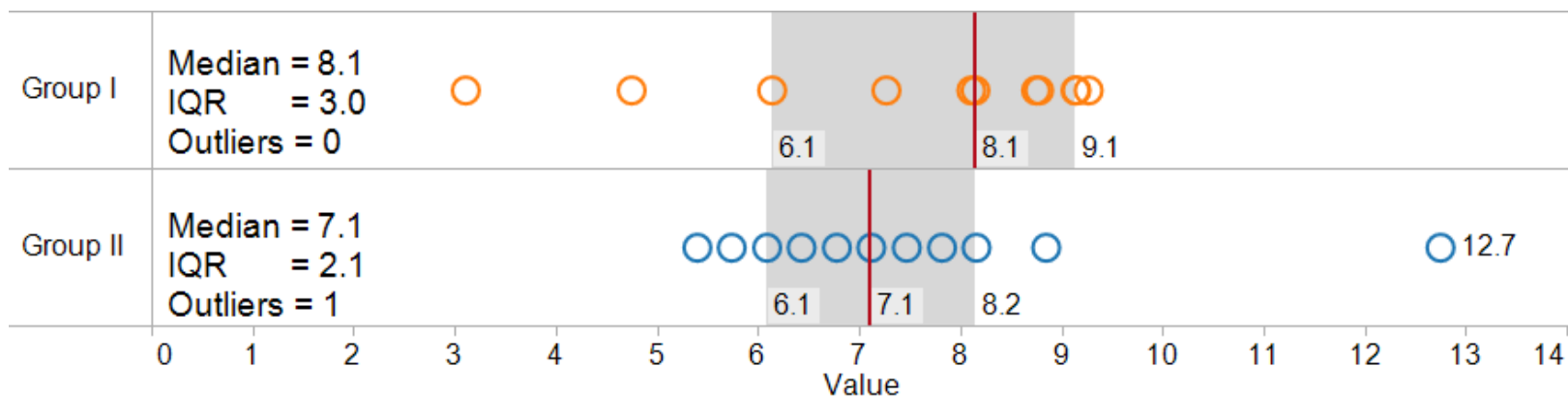
dayche.com | گروه دایچه

# تحلیل آماری

## آمار توصیفی برای متغیرهای کمی

### □ دیگر شاخص های پراکندگی

دامنه میان چارکی: اختلاف چارک اول و سوم داده ها



تولید محتوا: زهرا ذوالقدر

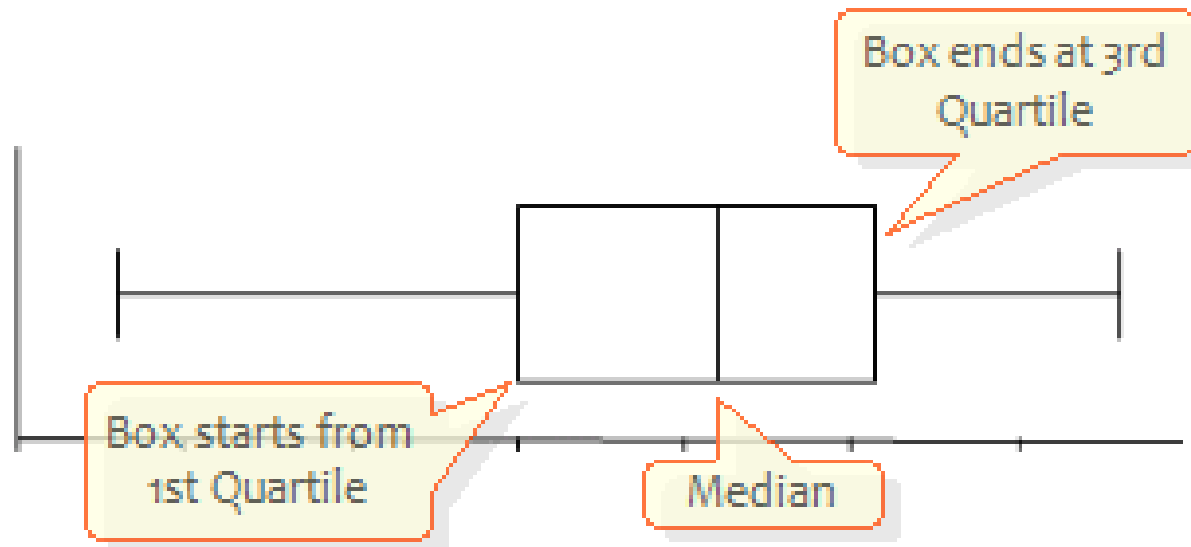
daychegroup

daychegroup

dayche.com | گروه دایچه

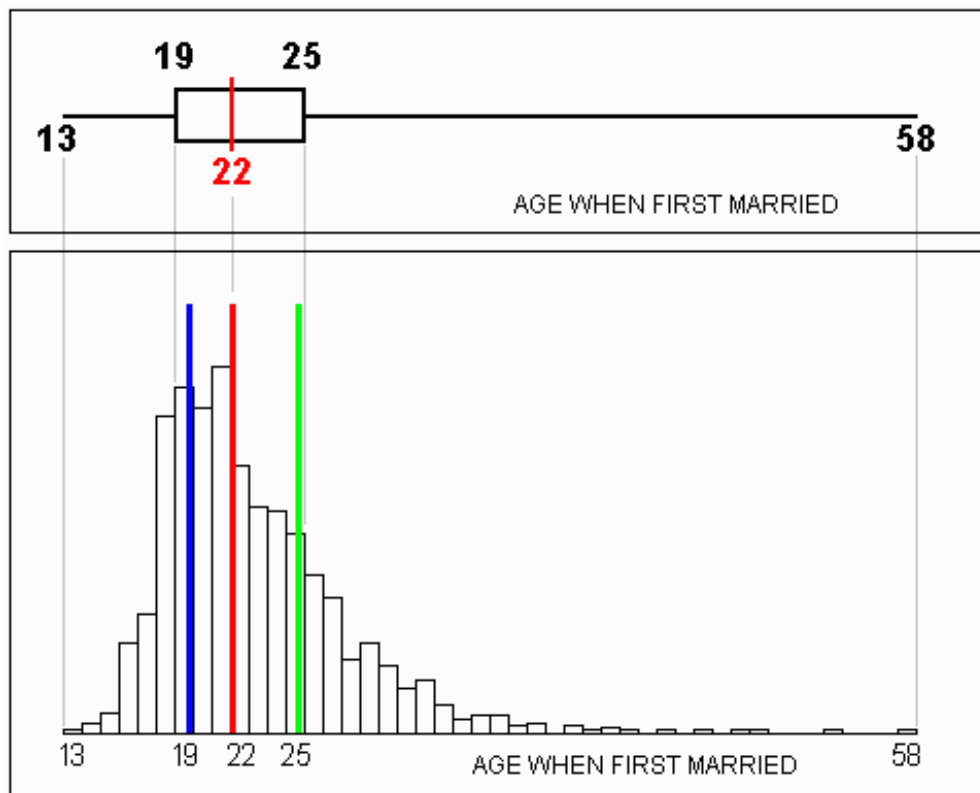
### □ دیگر شاخص های پراکندگی

دامنه میان چارکی: اختلاف چارک اول و سوم داده ها- نمودار جعبه ای



### □ دیگر شاخص های پراکندگی

دامنه میان چارکی: اختلاف چارک اول و سوم داده ها- نمودار جعبه ای



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تحلیل آماری

## آمار استنباطی



برآورد نقطه ای


برآورد فاصله ای

آزمون فرض

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایکه 

### برآورد نقطه ای □

هدف به دست آوردن یک مقدار عددی از داده هاست که تا حد ممکن به مقدار نامعلوم پارامتر نزدیک باشد.


## Point Estimates

Population Parameter	Unbiased estimator	Formula
Mean, $\mu$	$\bar{x}$	$\bar{x} = \frac{\sum x_i}{n}$
Variance, $\sigma^2$	$s^2$	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$
Proportion, $\pi$	$p$	$p = \frac{x \text{ successes}}{n \text{ trials}}$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 

# تحلیل آماری

## آمار استنباطی



### □ برآورد نقطه ای

ویژگی های برآوردگر:

○ ناریبی


○ کمترین واریانس

○ سازگاری

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

~ گروه دایکه | dayche.com 

### برآورد نقطه ای □

دقت برآورد نقطه ای: خطای استاندارد

The formula to find the sample mean

$$\mu_x = \frac{\sum_{i=1}^n x_i}{n}$$

Formula to estimate sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu_x)^2}{n - 1}}$$


Formula to estimate **standard error (SE) of mean**

$$SE_{\mu_x} = \frac{s}{\sqrt{n}}$$

تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

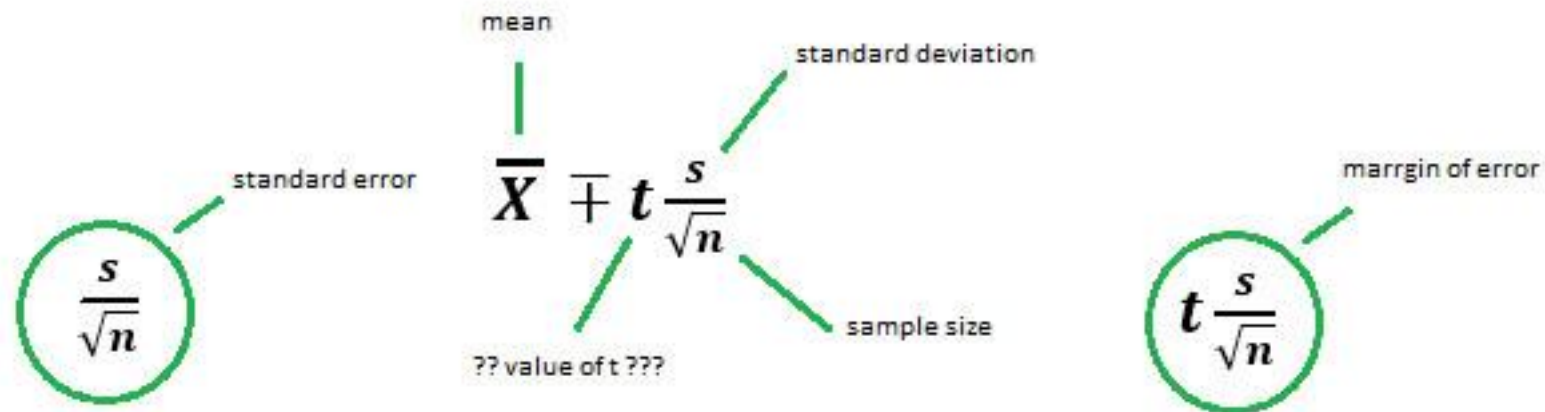
dayche.com | گروه دایچه 



# تحليل آماری

## آمار استنباطی


فاصله اطمینان □

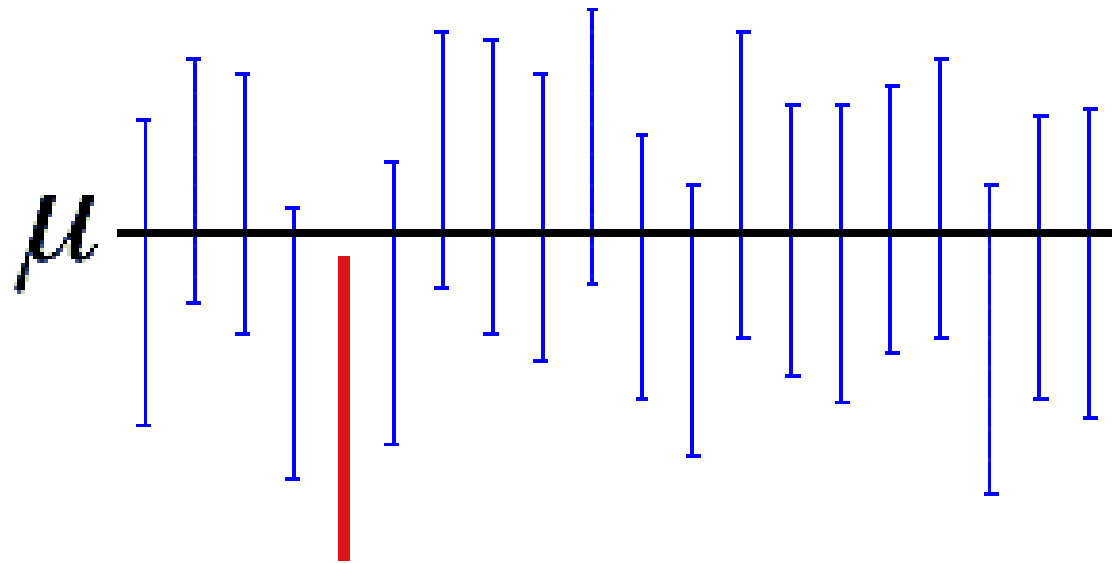


تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

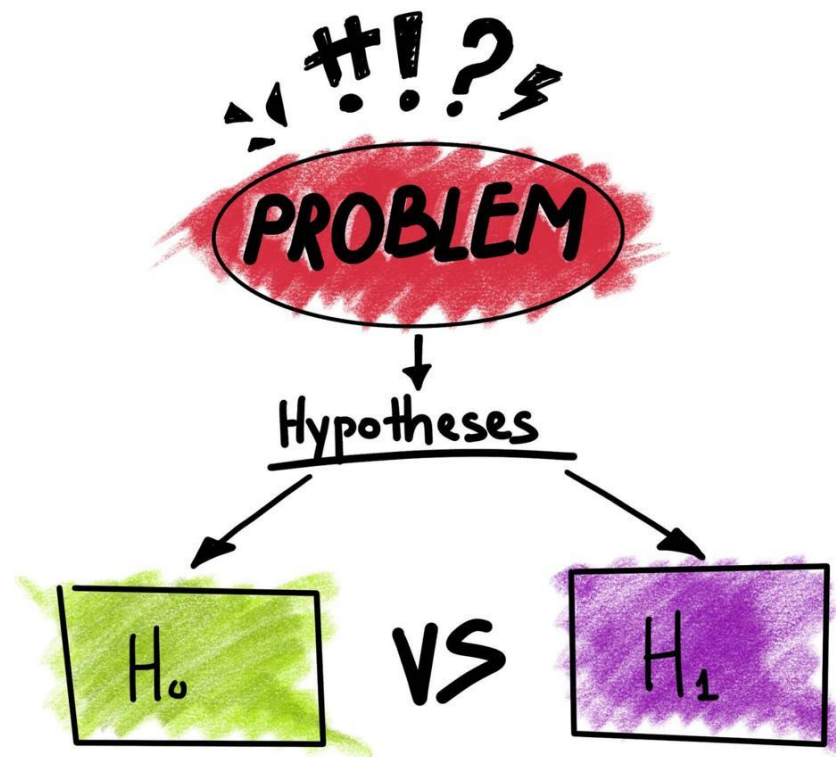
dayche.com | گروه دایچه 



A 95% confidence interval indicates that 19 out of 20 samples (95%) from the same population will produce confidence intervals that contain the population parameter.

### آزمون فرض □

فرض آماری: یک ادعا یا یک گزاره در مورد جمعیت یا در مورد پارامتر توزیع جمعیت



تولید محتوا: زهرا ذوالقدر

daychegroup 

daychegroup 

dayche.com | گروه دایچه 