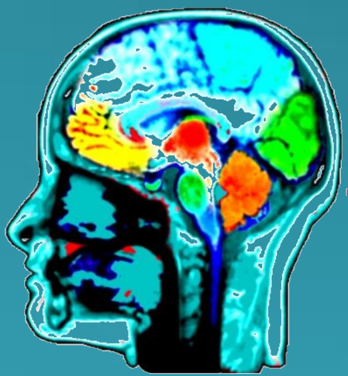
# Introduction To Data Mining

Isfahan University of Technology (IUT)

Bahman 1401



## Review of Probability Theory

Dr. Hamidreza Hakim
hamid.hakim.u@gmail.com

# Content

- Elements of Probability
- Random variables
- Two random variables
- Bayes Rule

Based on ”Review of Probability Theory” from CS 229 Machine Learning, Stanford University (Handout posted on the course website)

# REVIEW OF PROBABILITY THEORY

# Elements of Probability

احتمال یه تابع است که ۳ تاویژگی زیر را داشته باشه:تابعی که بتونه فضای حالت هامون را به مقداری ببره که همیشه مثبت است

- Sample space $\Omega$: the set of all the outcomes of an experiment

فضای نمونه ای مثلا برای قد انسان ها میشه یه متغیر پیوسته

- Event space $F$: a collection of possible outcomes of an experiment. $F \subseteq \Omega$.
- Probability measure: a function $P$: $F \rightarrow R$ that satisfies the following properties:

انداختن تاس، ۶تا حالت داره و گسسته است و سکه انداختن دوتا حالت داره پس همه ی حالت هایی که ازمایش ما میتونه داشته باشه میشه فضای نمونه ای

فضای پیشامد میشه گفت یه زیرمجموعه ای از فضای نمونه است مثلا بازه ی خاصی از قد انسان ها

  - $P(A) \geq 0 \ \forall \ A \in F$
  - $P(\Omega) = 1$

گه همه ی اعضای مجموعه مون ی ورودی اون تابع باشه، مقدار یک را به ما بده

  - If $A_1, A_2, \ldots$ are disjoint events, then

اگه راجع به تکه های مختلف مجموعمون حرف بزنیم مجموع احتمال اونها ا احتمال مجموع اونها یکی بشه

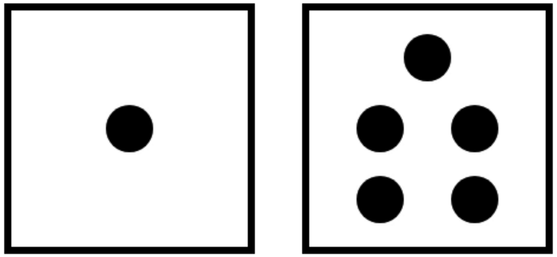
$$\boxed{P(\cup_i A_i) = \sum_i P(A_i)}$$

احتمال مجموع

مجموع احتمال ها

# Elements of Probability(Example)

- tossing a six-sided die
- Measure human Color



# Properties of Probability

قوانین و ویژگی های تابع احتمال

- If $\boxed{A \subseteq B} \implies \boxed{P(A) \leq P(B)}$
- $\boxed{P(A \cap B) \leq \min\,(P(A), P(B))}$
- $\boxed{P(A \cup B) \leq P(A) + P(B)}$ (Union Bound)
- $\boxed{P(\Omega \setminus A) = 1 - P(A)}$

مکمل زمان هایی که تاس ۲ میاد مثلا

مکمل فضا

- If $A_1, \ldots, A_k$ is a disjoint partition of $\Omega$, then

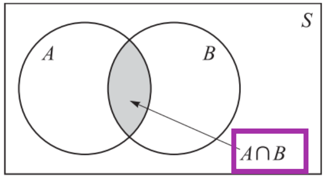
$$\boxed{\sum_{i=1}^{k} P(A_k) = 1}$$

# Conditional Probability

احتمال شرطی

- A conditional probability $P(A|B)$ measures the probability of an event A after observing the occurrence of event $B$

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




- Two events $A$ and $B$ are independent iff $P(A|B) = P(A)$ or equivalently,

دو پیشامد نسبت به هم مستقل هستند

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

# Conditional Probability(Examples)

- A math teacher gave her class two tests. 25% of the class passed both tests and 42% of the class passed the first test. What percent of those who passed the first test also passed the second test?
- In New England, 84% of the houses have a garage and 65% of the houses have a garage and a back yard. What is the probability that a house has a backyard given that it has a garage?

# Independent Events Examples

- What's the probability of getting a sequence of 1,2,3,4,5,6 if we roll a dice six times?

- A school survey found that 9 out of 10 students like pizza. If three students are chosen at random with replacement, what is the probability that all three students like pizza?

# Random Variable

متغیرهای تصادفی یه مجموعه ای از مقادیر را میتوانند بگیرند

A random variable $X$ is a function that maps a sample space $\Omega$ to real values. Formally,

$$\boxed{X : \Omega \longrightarrow R}$$

اگه یه سکه را ۱۰ بار انداختم چند بارش رو
میاد چند بار پشت؟ اینجا داریم درباره ی یه
چیز کمی صحبت میکنیم که مقدار داره
در این زمان ها سراغ متغیرهای تصادفی
میریم
مثال : تعداد دفعاتی که یک سکه رو میاد
برای ۱۰بار انداختن
یا تعداد دفعاتی که تاس زوج میاد در ۶بار
انداختن
تعداد دانش اموزانی که قدشان از یه حدی
بیشتر است

Examples:

- Rolling one dice
  $X$ = number on the dice at each roll
- Rolling two dice at the same time
  $X$ = sum of the two numbers

درباره ی اون مقداری که قابل اندازه
گیری است صحبت کنیم

# Random Variable

A random variable can be continuous. E.g.,

- $X$ = the length of a randomly selected phone call
  (What's the $\Omega$?)

تعریف طول شماره تلفن ها به
عنوان متغیرتصادفی
ثلا طول شماره تلفن های همراه و
ثابت متفاوت است پس چندتا حالت
داریم

- $X$ = amount of coke left in a can marked 12oz
  (What's the $\Omega$?)

# Probability Mass Function

If $X$ is a discrete random variable, we can specify a probability for each of its possible values using the probability mass function $(PMF)$. Formally, a $PMF$ is a function $p\colon \Omega \longrightarrow R$ such that

میخایم پدیده ی تصادفی را بیشتر بشناسیم
اگه متغیرتصادفی همیشه مقادیر گسسته داشته باشه همیشه یه پی ام اف هم خواهد داشت

احتمال رخدادن اینکه تاس مثلا عدد یک بیاد یک ششم است که بش پی ام اف میگیم

$$\boxed{p(x) = P(X = x)}$$

- Rolling a dice:
  $p(X = i) = \frac{1}{6} \quad i = 1, 2, \ldots, 6$

متغیرتصادفی: مقادیری که به ازای انداختن تاس بدست میاریم باشه

- Rolling two dice at the same time:
  $X$ = sum of the two numbers
  $p(X = 2) = \frac{1}{36}$

pmf:
احتمال رخدادن یه حالتی ازمتغیرتصادفی را میگه

اگه متغیرتصادفی مان را جمع مقادیری بگیریم که دوتاس در دوبار پرتاب میگیرن

A Probability Mass Function (PMF) is a function that describes the probability distribution of a discrete random variable. It assigns a probability to each possible value that the random variable can take on. Here are some examples of PMFs:

Fair Coin Toss: Suppose we toss a fair coin once and let X be the number of heads that come up. The possible values for X are 0 and 1, and the PMF is given by:

P(X=0) = 1/2
P(X=1) = 1/2

This PMF tells us that there is a 50% chance of getting zero heads and a 50% chance of getting one head when we toss a fair coin once.

Rolling a Die: Suppose we roll a six-sided die once and let Y be the number that comes up. The possible values for Y are 1, 2, 3, 4, 5, and 6, and the PMF is given by:

P(Y=1) = 1/6
P(Y=2) = 1/6
P(Y=3) = 1/6
P(Y=4) = 1/6
P(Y=5) = 1/6
P(Y=6) = 1/6

This PMF tells us that each of the six possible outcomes has an equal chance of occurring when we roll a fair die once.

Drawing Marbles from a Bag: Suppose we have a bag with three red marbles and two blue marbles. We draw one marble from the bag without replacement and let Z be the color of the marble we draw (where R represents red and B represents blue). The possible values for Z are R and B, and the PMF is given by:

P(Z=R) = 3/5
P(Z=B) = 2/5

This PMF tells us that there is a 60% chance of drawing a red marble and a 40% chance of drawing a blue marble when we draw one marble from the bag without replacement.

These are just a few examples of PMFs. In general, the PMF can be used to describe the probability distribution of any discrete random variable.

To identify the random variable in a probability problem, you need to look for the quantity that can take on different values with some probability. A random variable is a mathematical representation of this quantity. Here are some steps to follow to identify the random variable in a probability problem:

Read the problem statement carefully and identify what is being measured or observed.

Determine which of these quantities varies from trial to trial or observation to observation.

Assign a symbol to represent the quantity identified in step 2. This symbol should be capitalized to denote that it represents a random variable (e.g., X, Y, Z).

Define the range of possible values that the random variable can take on. This range should include all possible values that the quantity can take on, even if some values have zero probability.

Determine the probability distribution of the random variable. This involves assigning probabilities to each possible value that the random variable can take on.

For example, consider the following problem: A fair six-sided die is rolled. What is the probability of rolling an even number?

In this problem, the quantity being measured is the number rolled on the die. The quantity varies from trial to trial, so it is a random variable. We can represent this random variable with the symbol X.

The range of possible values for X is {1, 2, 3, 4, 5, 6}. Since the die is fair, each of these values has probability 1/6.

The probability distribution of X is given by:

P(X=1) = 1/6
P(X=2) = 1/6
P(X=3) = 1/6
P(X=4) = 1/6
P(X=5) = 1/6
P(X=6) = 1/6

Therefore, we can say that X is a discrete random variable with a uniform distribution.

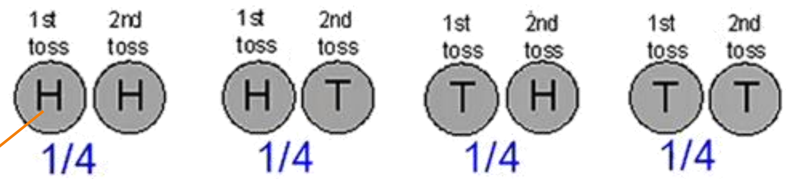
In summary, to identify the random variable in a probability problem, you need to look for the quantity that can take on different values with some probability and assign a symbol to represent it.

# Probability Mass Function(Examples)

- X: Number of Tail

متغیرتصادفی مان اینجا
عداد پشت یا تیل هایی که
سکه میاد

سکه ی اول یا سکه ی
دوم پشت بیاد دوتا حالت
میشه




pmf
یه تابعی از جنس احتمال است
احتمال رخداد یه متغیر تصادفی
را گزارش میکنه

X=0

X=1

X=2

نوشتن حالت های متغیرتصادفی
و تعداد رخدادهای اون حالت از
متغیرتصادفی را مینویسیم توش

List of possible values

Probability of each value

| X | 0 | 1 | 2 |
|---|---|---|---|
| P(X=x) | 1/4 | 1/2 | 1/4 |

# Probability Mass Function(Examples)

- X be the number of tails in Flipping a Coin Three Times

| Outcome | Probability | X |
|---|---|---|
| HHH | 1/2*1/2*1/2=1/8 | 0 |
| HHT | 1/8 | 1 |
| HTH | 1/8 | 1 |
| THH | 1/8 | 1 |
| HTT | 1/8 | 2 |
| THT | 1/8 | 2 |
| TTH | 1/8 | 2 |
| TTT | 1/8 | 3 |

# Probability Mass Function(Examples)

- X be the number of tails in Flipping a Coin Three Times

| Outcome | Probability | X |
|---|---|---|
| HHH | 1/8 | 0 |
| HHT | 1/8 | 1 |
| HTH | 1/8 | 1 |
| THH | 1/8 | 1 |
| HTT | 1/8 | 2 |
| THT | 1/8 | 2 |
| TTH | 1/8 | 2 |
| TTT | 1/8 | 3 |

$0 \rightarrow \boxed{1/8}$

$1 \rightarrow 1/8+1/8+1/8= \boxed{3/8}$

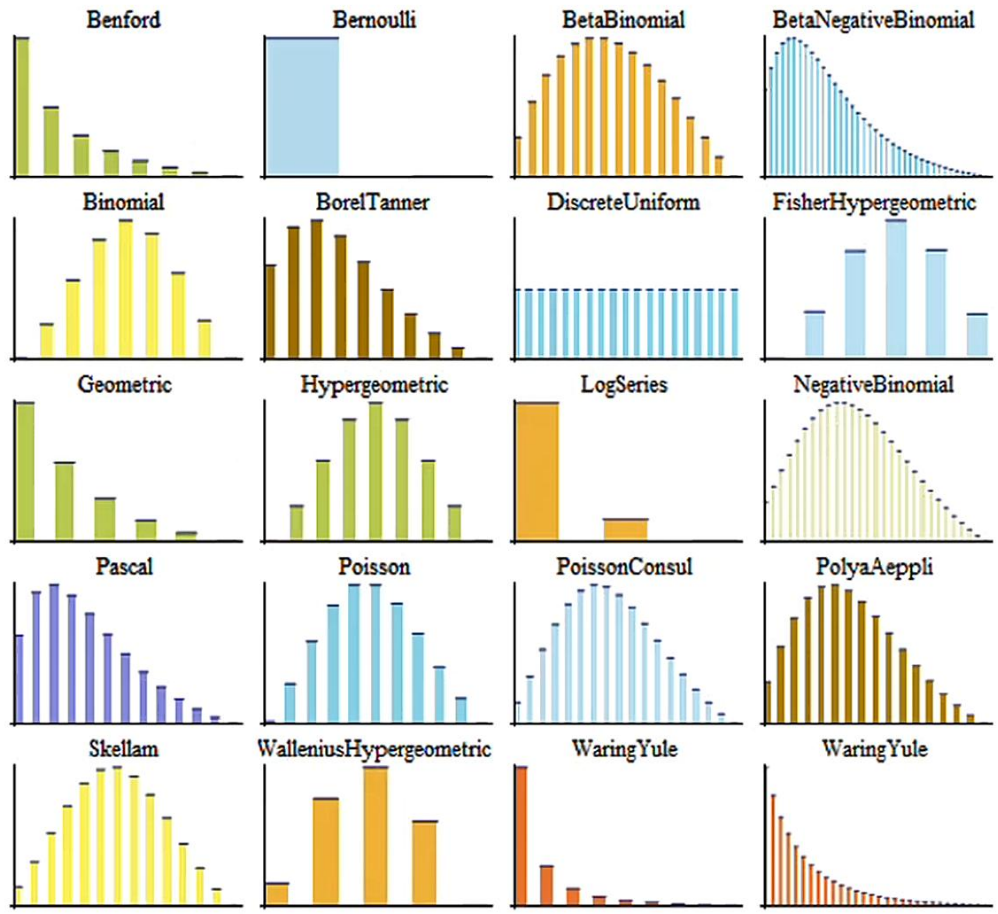
$2 \rightarrow 1/8+1/8+1/8= \boxed{3/8}$

$3 \rightarrow \boxed{1/8}$

# Probability Mass Function

| X | $x_1$ | $x_2$ | $x_3$ | … | $x_n$ |
|---|---|---|---|---|---|
| P(X=x) | $p_1$ | $p_2$ | $p_3$ | … | $p_n$ |

محور ایکس مقادیر متغیرتصادفی است
و محور وای احتمال رخداد اون متغیر
است



# Probability Mass Function

- $\boxed{X \sim Bernoulli(p),\ p \in [0,1]}$

$$p(x) = \begin{cases} p & \text{if} \quad x = 1 \\ 1-p & \text{if} \quad x = 0 \end{cases}$$

- $\boxed{X \sim Binomial(n,p),\ p \in [0,1] \text{ and } n \in Z^+}$

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

- $\boxed{X \sim Geometric(p),\ p > 0}$

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

- $\boxed{X \sim Poisson(\lambda),\ \lambda > 0}$

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

متغیر تصادفی پیوسته

# Probability Density Function

- If $X$ is a continuous random variable, we can NOT specify a probability for each of its possible values (why?)
- We use a probability density function $PDF$ to describe the relative likelihood for a random variable to take on a given value
- A ($PDF$) specifies the probability of $X$ takes a value within a range. Formally, a $PDF$ is a function $f(x)\colon \Omega \longrightarrow R$ such that

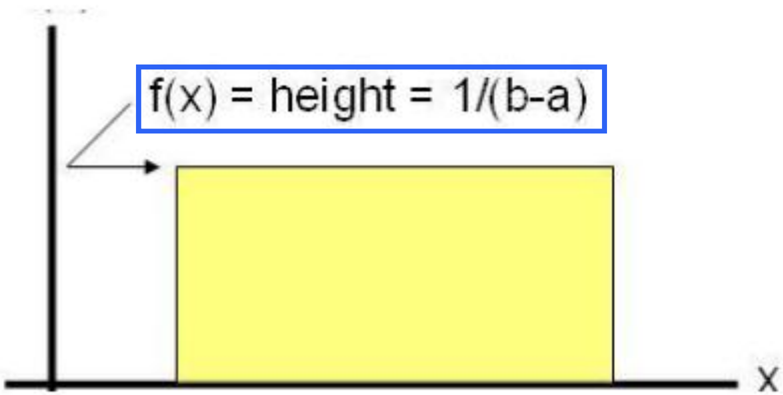
درفضای پیوسته ما راجع به مقدار دقیق یه عدد
حرف نمیزنیم بحث های حد داره
مثلا اینکه احتمال اینکه قد بین ۱۶۰ تا ۱۸۰ باشه
چقدره؟ باید یه بازه بش بدیم

$$P(a < X < b) = \int_a^b f(x)dx$$

# Probability Density Function

- $X \sim$ uniform on $[a, b]$:

توزیع به صورت یکنواخت است یعنی همه ی حالت های متغیر تصادفی شانس برابری دارند ینی احتمال تک تک این مقادیر یکسان است



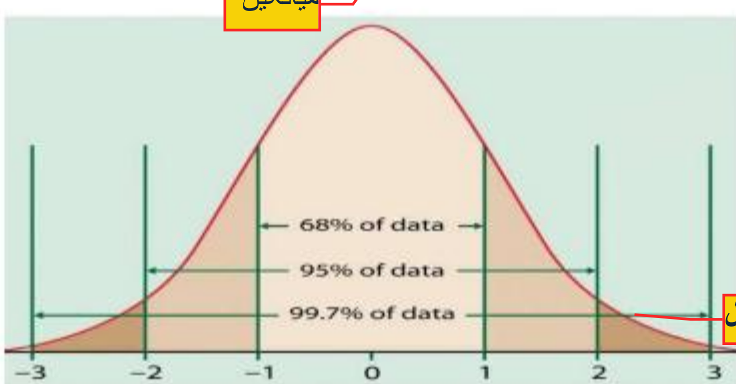

$$\boxed{f(x) = \frac{1}{b-a}}$$

فقط ابتدا و انتهای بازه را بش میدیم خودش ارتفاع را میده

- $X \sim \boxed{N(\mu, \sigma)}$:

انحراف معیار

میانگین




توزیع نرمال

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

# Expected value of Random Variable

مقدار امیدریاضی یک متغیر یا
مقدار مورد انتظار ما

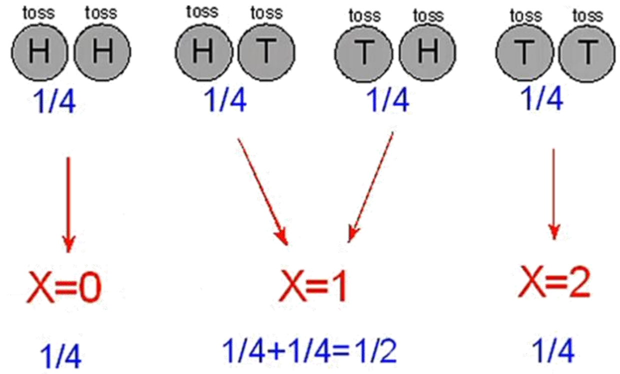
از جنس احتمال نیست!!!
از جنس همون متغیریه
که داریم دربارش حرف
میزنیم

$$E(X) = \sum_{x} x f(x)$$

ضرب احتمال در مقادیر
رخداد و جمع اینها باهم

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

مقدار متوسط پشت اومدن دوتا سکه؟




List of possible values

Probability of each value

| x | 0 | 1 | 2 |
|---|---|---|---|
| P(X=x) | 1/4 | 1/2 | 1/4 |

# Variance of Random Variable

واریانس یه متغیر تصادفی
میشه پراکندگی اون متغیر

$$\boxed{Var(X)} = \boxed{\sigma^2} = \boxed{E\left[(X-\mu)^2\right]}$$

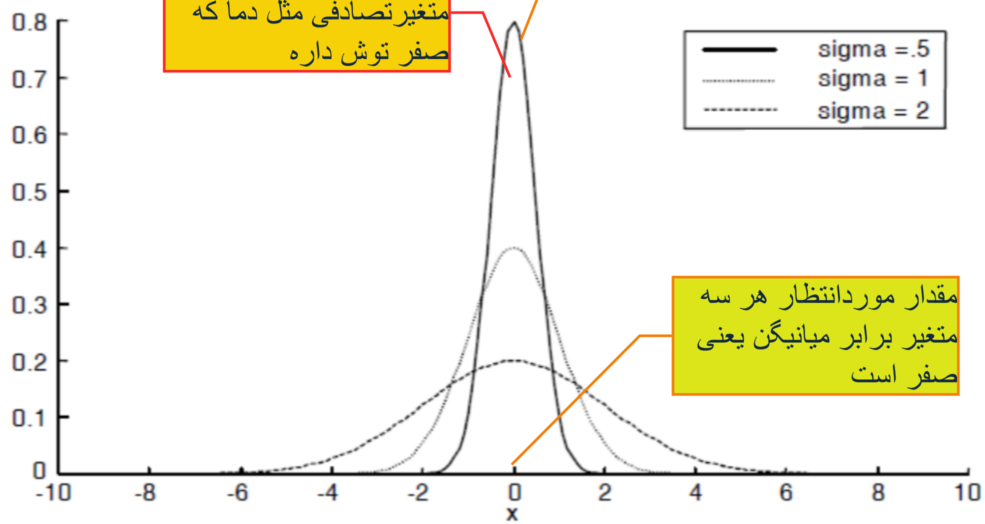
مقادیر متغیر تصادفی چه
طیفی داره؟
قدار های متغیرتصادفی را
نهای میانگینشون میکنه و
بعد امیدریاضیشون را
حساب میکنه

توزیع نرمال

- σ = Deviation Standard

صفر خیلی توش رخداده

پی دی اف یک
متغیرتصادفی مثل دما که
صفر توش داره




مقدار موردانتظار هر سه
متغیر برابر میانیگن یعنی
صفر است

# More Than One Random Variable(Example)

صحبت درباره ی قد و وزن ادم ها دراینجا دوتا متغیر تصادفی داریم

یک پدیده ولی دوتا بعد داره

- Flip a coin ten times
- $X(\omega)$ = the number of heads that come up as well as
- $Y(\omega)$ = the length of the longest run of consecutive heads

طول طولانی ترین سرهای متوالی

# Joint Probability Mass Function

وقتی چندتا متغیرتصادفی
داریم یه تابع توزیع تعریف
میکنیم که بهش تابع توزیع
توام اون متغیر ها میگیم ینی
دوست داریم راجع به
دوتاشون اطلاعات کسب کنیم

چرا میگیم
mass function
چون متغیر هامون گسسته هستند

If we have two discrete random variables $X, Y$, we can define their joint probability mass function (PMF) $p_{XY} : R^2 \longrightarrow [0, 1]$ as:

چه تعداد ادم داریم که این
وزن و قد را داشته باشند؟

$$p(x, y) = P(X = x, Y = y)$$

where $p(x, y) \leq 1$ and $\sum_{x \in X} \sum_{y \in Y} p(x, y) = 1$

تابع توزیع توام متغیرهامون که بهش
joint probability mass function
میگیم

- $X, Y$: rolling two dice
  $p(x, y) = \frac{1}{36} \quad x, y = 1, 2, \ldots, 6$

- $X$: rolling one dice $\quad Y$: drawing a colored ball
  $p(6, green) = ? \quad p(5, red) = ?$

انداختن دوتا تاس
تاس شماره یک میشه متغیر تصادفی اول
تاس شماره دو میشه متغیر دوم
وضعیت این دوتا تاس باهمدیگر و احتمال رخدادنشون را
میشه به کمک این تابع بدست اورد

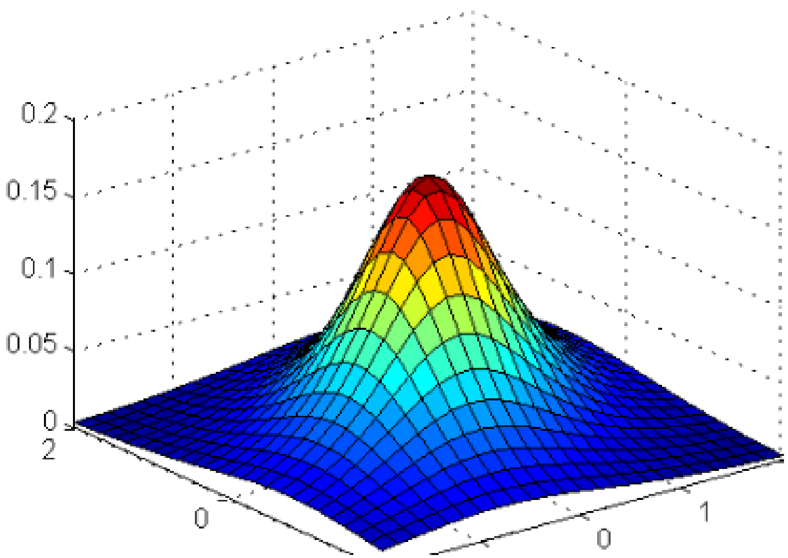
# Joint Probability Density Function

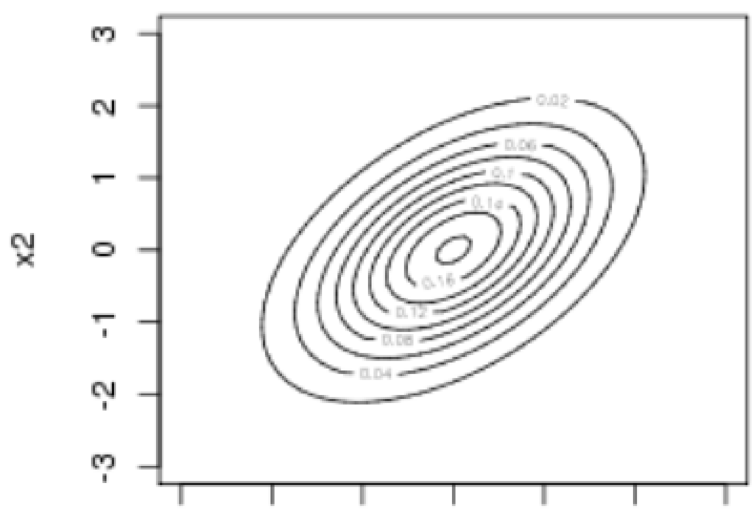
If we have two continuous random variables $X, Y$, we can define their joint probability density function $(PDF)$ $\boxed{f_{XY}} : R^2 \longrightarrow [0,1]$ as:

$$P(a < \boxed{X} < b, c < \boxed{Y} < d) = \boxed{\int_c^d \int_a^b f(x,y)dxdy}$$

- 2D Gaussian

فضای پیوسته



# Marginal Probability Mass Function

How does the joint *PMF* over two discrete variables relate to the *PMF* for each variable separately? It turns out that

توزیع حاشیه ای پیداکنیم

$$\boxed{p(x) = \sum_{y \in Y} p(x, y)}$$

پیداکردن احتمال یک متغیر تصادفی از احتمال توام اون متغیر با بقیه پیدا کنیم باید روی بعدی که نمیخایم یه جمع انجام بدیم
مثلا p(x,y) را داریم میخایم راجع به (p(x حرف بزنیم

- $X, Y$: rolling two dice

$$p(x, y) = \frac{1}{36} \quad x, y = 1, 2, \ldots, 6$$

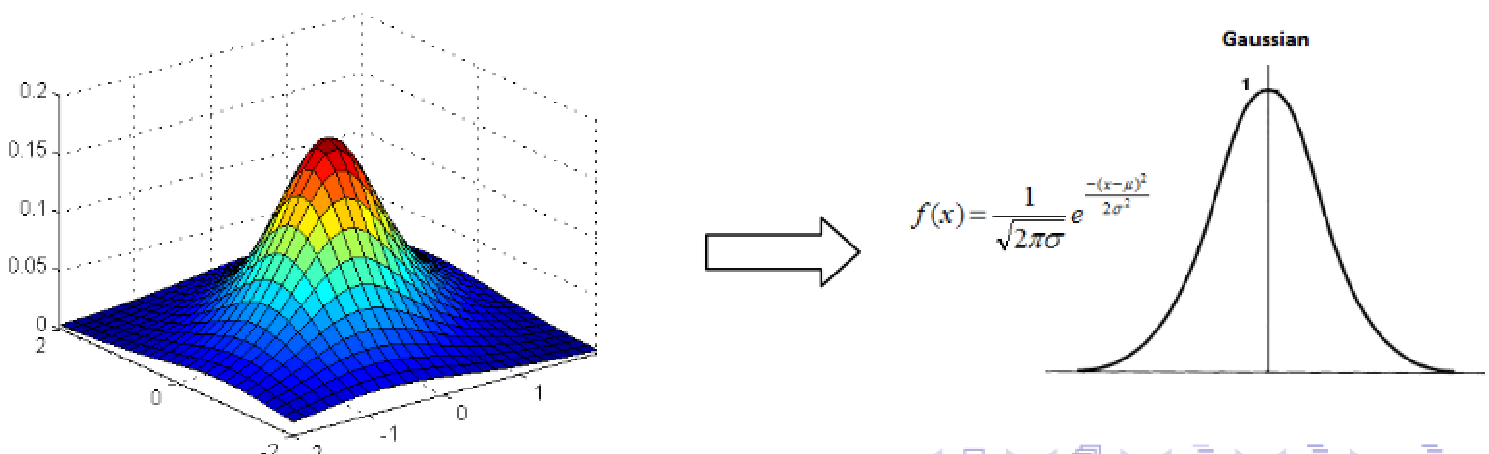
$$\boxed{p(x) = \sum_{y=1}^{6} p(x, y) = \frac{1}{6}}$$

# Marginal Probability Density Function

Similarly, we can obtain a marginal *PDF* (also called marginal density) for a continuous random variable from a joint *PDF*:

$$\boxed{f(x) = \int_{-\infty}^{\infty} f(x, y)dy}$$

- Integrating out one variable in the 2D Gaussian gives a 1D Gaussian in either dimension



# Conditional Probability Distribution

A conditional probability distribution defines the probability distribution over $Y$ when we know that $X$ must take on a certain value $x$

- Discrete case: conditional $PMF$

احتمال توام ایکس و وای

$$p(y|x) = \frac{p(x,y)}{p(x)} \Longleftrightarrow \boxed{p(x,y) = p(y|x)p(x)}$$

- Continuous case: conditional $PDF$

$$f(y|x) = \frac{f(x,y)}{f(x)} \Longleftrightarrow \boxed{f(x,y) = f(y|x)f(x)}$$

# Marginal vs. Conditional

متغیرهای تصادفی مون اینجا انداختن دوتا تاس است
احتمال اینکه هردو یک بیاد یک تقسیم بر ۳۶ است

- **Marginal probability:**

احتمال اینکه هر دو یک بیاد

احتمال اینکه تاس اول ۴ بیاد میشه مارجینال

| $i \backslash j$ | 1 | 2 | 3 | 4 | 5 | 6 | $p_X(i)$ |
|---|---|---|---|---|---|---|---|
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| $p_Y(j)$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | |

احتمال اینکه تاس اول ۴ بیاد میشه ۱/۶

- **Conditional probability**: probability of rolling a 2

احتمال اینکه تاس دوم ۲ یاد به شرط ۲ بودن تاس اول

| $i \backslash j$ | 1 | 2 | 3 | 4 | 5 | 6 | $p_X(i)$ |
|---|---|---|---|---|---|---|---|
| 1 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 2 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 3 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 4 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 5 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| 6 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/36 | 1/6 |
| $p_Y(j)$ | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | |

# Bayes Rule

قوانین بیز
مشکل اصلی در احتمال ها
محاسبه ی مدل توزیعی است
p(x,y

- We can express the joint probability in two ways:

$$p(x, y) = p(y|x)p(x)$$
$$p(x, y) = p(x|y)p(y)$$

ساخت توزیع توام به کمک احتمال شرطی و احتمال تکی

- Bayes rule:

حساب کردن توزیع شرطی

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \quad \text{(discrete)}$$

جنسش احتمال نیست

$$f(y|x) = \frac{f(x|y)f(y)}{f(x)} \quad \text{(continuous)}$$

$$P(Y) = \sum_i P(Y|X_i)P(X_i)$$

Proof: law of total probability:

پیداکردن توزیع یا احتمال یک متغیر

قانون احتمال کل

# Bayes Rule Application

A patient underwent a HIV test and got a positive result. Suppose we know that

- Overall risk of having HIV in the population is 0.1%
- The test can accurately identify 98% of HIV infected patients
- The test can accurately identify 99% of healthy patients

What's the probability the person indeed infected HIV?

Bayes' rule is a fundamental theorem in probability theory that describes the probability of an event based on prior knowledge of related events. Here are some examples of Bayes' rule:

Medical Diagnosis: Suppose a patient undergoes a medical test for a disease that has a prevalence rate of 5 % in the population. The test has a false positive rate of 10% and a false negative rate of 5%. If the patient tests positive, what is the probability that he/she actually has the disease? Using Bayes' rule, we can calculate the probability as follows:

P(Disease|Positive Test) = P(Positive Test|Disease) * P(Disease) / P(Positive Test)

where,
P(Positive Test|Disease) = 95% (True Positive Rate)
P(Disease) = 5%
P(Positive Test) = P(Positive Test|Disease) * P(Disease) + P(Positive Test|No Disease) * P(No Disease)
= 95% * 5% + 10% * 95%
= 9.25%

Therefore,
P(Disease|Positive Test) = 95% * 5% / 9.25% = 51.35%

Hence, the probability that the patient actually has the disease given a positive test result is about 51.35%.

Email Spam Filtering: Suppose an email filtering system receives a new email with certain words that occur frequently in spam emails. If 40% of all emails are spam, and the occurrence of these words in spam emails is 70%, while the occurrence of these words in non-spam emails is 20%, what is the probability that this email is spam? Using Bayes' rule, we can calculate the probability as follows:

P(Spam|Words) = P(Words|Spam) * P(Spam) / P(Words)

where,
P(Words|Spam) = 70%
P(Spam) = 40%
P(Words) = P(Words|Spam) * P(Spam) + P(Words|No Spam) * P(No Spam)
= 70% * 40% + 20% * 60%
= 38%

Therefore,
P(Spam|Words) = 70% * 40% / 38% = 73.68%

Hence, the probability that this email is spam given the occurrence of certain words is about 73.68%.

Sure, here are some numeric examples of Bayes' rule that are slightly more challenging:

Drug Testing: A drug test is 99% accurate in detecting a banned substance in an athlete's urine sample. However, the probability of a healthy, non-doping athlete testing positive on the test is 0.1%. If an athlete tests positive, what is the probability that they actually used the banned substance?

P(Doping|Positive Test) = P(Positive Test|Doping) * P(Doping) / P(Positive Test)

where,
P(Positive Test|Doping) = 99%
P(Doping) = 0.5%
P(Positive Test) = P(Positive Test|Doping) * P(Doping) + P(Positive Test|No Doping) * P(No Doping)
= 99% * 0.5% + 0.1% * 99.5%
= 0.598%

Therefore,
P(Doping|Positive Test) = 99% * 0.5% / 0.598% = 82.78%

Hence, if an athlete tests positive, there is still an 82.78% chance they actually used the banned substance.

Disease Testing with Multiple Symptons: Suppose a particular disease has a prevalence rate of 1% in a population. The disease causes two symptoms, A and B, with probabilities as follows: P(A|Disease) = 80%, P(B|Disease) = 70%, P(A|No Disease) = 10%, and P(B|No Disease) = 20%. If a person has both symptoms A and B, what is the probability that they have the disease?

P(Disease|A,B) = P(A,B|Disease) * P(Disease) / P(A,B)

where,
P(A,B|Disease) = P(A|Disease) * P(B|Disease) = 80% * 70% = 56%
P(Disease) = 1%
P(A,B) = P(A,B|Disease) * P(Disease) + P(A,B|No Disease) * P(No Disease)
= 56% * 1% + 10% * 20% * 99%
= 11.84%

Therefore,
P(Disease|A,B) = 56% * 1% / 11.84% = 4.72%

Hence, if a person has both symptoms A and B, there is only a 4.72% chance that they actually have the disease.

These examples illustrate how Bayes' rule can be used to calculate probabilities when dealing with more complex scenarios like drug testing or diseases with multiple symptoms.

# Bayes Rule - Application

We have two random variables here:

- $X \in \{+, -\}$: the outcome of the HIV test
- $C \in \{\mathrm{Y}, \mathrm{N}\}$: the patient has HIV or not

We want to know: $P(C{=}\mathrm{Y}|X{=}+)$?

Apply Bayes rule:

p(x =+) = p(x =+ | c =y) * p(c =y) + p(x=+ | c= N) * p(c =N)
p(x = +) = 0.98 * 0.001 + (1 - 0.99 ) * (1 - 0.001) =

$$P(C{=}\mathrm{Y}|X{=}+) = \frac{P(X{=}+|C{=}\mathrm{Y})P(C{=}\mathrm{Y})}{P(X{=}+)}$$

$P(X{=}+|C{=}\mathrm{Y}) = 0.98$ $\qquad$ $P(C{=}\mathrm{Y}) = 0.001$

p(x = + | C=N) = 1- p(x =- | C=N)

$$P(X{=}+) = 0.98 * 0.001 + (1\text{-}0.99) * 0.999 = 0.01097$$

p(C = N) = 1-0.001 = 0.999

Answer: $0.98 * 0.001/0.01097 = 8.9\%$

# Independence

Two random variables X and Y are independent iff

- For discrete random variables

  $$\boxed{p(x, y) = p(x)p(y) \quad \forall x \in X, y \in Y}$$

- For discrete random variables

  $$\boxed{p(y|x) = p(y)} \qquad \forall y \in Y \text{ and } p(x) \neq 0$$

- For continuous random variables

  $$\boxed{f(x, y) = f(x)f(y) \quad \forall x, y \in R}$$

- For continuous random variables

  $$\boxed{f(y|x) = f(y)} \qquad \forall y \in R \text{ and } f(x) \neq 0$$

# Multiple Random Variables

پیداکردن توزیع توام ان تا متغیر تصادفی

Extend to multiple random variables :

- Joint Distribution (discrete):

$$p(x_1, \dots, x_n) = P(X1 = x_1, \dots, X_n = x_n)$$

- Conditional Distribution (chain rule - discrete)

$$\boxed{p(x_1, \dots, x_n)} = p(x_n|x_1, \dots, x_{n-1})p(x_1, \dots, x_{n-1})$$

$$= p(x_n|x_1, \dots, x_{n-1})p(x_{n-1}|x_1, \dots, x_{n-2})p(x_1, \dots, x_{n-2})$$

$$= \boxed{p(x_1) \prod_{i=2}^{n} p(x_i|x_1, \dots, x_{i-1})}$$

نوشتن توزیع توام به کمک
مجموعه های توزیع شرطی

(continuous case can be defined similarly using *PDF*)

p(x1) * p(x2|x1) * p(x3|x1,x2) * p(x4 |x1,x2,x3) *....p(xn|x1,x2,...xn-1)

# Multiple Random Variables

- Independence:

  Discrete case: $X_1, \ldots, X_n$ are independent iff

  $$\boxed{p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i)}$$

  Continuous case: $X_1, \ldots, X_n$ are independent iff

  $$\boxed{f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i)}$$

# SOME OTHER POINTS

# Probabilistic View of a Dataset

What about a $\boxed{\text{dataset}}$ $S = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)\}$?

- We can view $S$ as $\boxed{d+1 \text{ random variables}}$ where $\boxed{d}$ is the number of $\boxed{\text{attributes in } \mathbf{x},}$ i.e.

$$X_1, \; X_2, \; \ldots, \; X_d, \; Y$$

- Uncover(model) $p(x_1, x_2, \ldots, x_d, y)$ from the training data

- For ANY $\boxed{(x_1, x_2, \ldots, x_n),}$ we will compute:

$$\boxed{P(y=0|x_1, x_2, \ldots, x_n)}?$$

$$\boxed{P(y=1|x_1, x_2, \ldots, x_n)}?$$

That is predicting $y$ from $\mathbf{x}$ !

# Bayes Rule Terminology

posterior

likelihood

prior

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

marginal probability

$\boxed{P(Y):}$ prior probability or, simply, $\boxed{\text{prior}}$

$\boxed{P(X|Y):}$ conditional probability or, $\boxed{\text{likelihood}}$

$\boxed{P(X):}$ marginal probability

$\boxed{P(Y|X)}$: posterior probability or, simply, $\boxed{\text{posterior}}$