Independence

Independence

Definition: Two events A and B are independent if $P(A \cap B) = P(A)P(B)$. Alternatively, they are independent if $P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Informally, two events are independent if they don't influence each other; knowing that A happened doesn't give you any additional information about B.

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

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

Conversely

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

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

Independence Example - Discrete

- Suppose that our sample space consists of N flips of a coin that has probability p of giving heads.
- The events corresponding to a H in position i and in position j are independent.
- The chance of getting k heads in N flips is

$$P(k, N) = \binom{N}{k} p^k (1-p)^{N-k}$$

The probability distribution on the set $0, \ldots, N$ given by this formula is called the *binomial distribution* for parameters p and N.

X = Sequences of H, T of length N

$= \{H, T\}^N$

$(H, T, H, H, \ldots, H) \in X$

$P(\text{Head in position } i) = p$ — the positions are independent.

$P(T \text{ " " " }) = 1-p$

If $E = (H, \ldots T, \ldots, H)$ — K heads, $N-K$ tails

$P(E) = p^K (1-p)^{N-K}$

~~Prob~~ Let $P(K, N)$ be the chance of K heads in N flips.

$P(K, N) = \binom{N}{K} p^K (1-p)^{N-K}$

Fix N: $P(K, N)$ = chance of K heads in N flips of a coin with $P(H) = p$.

$K = 0, 1, \ldots, N$

Binomial Distribution with parameters p, N.

Independence Example - Continuous

Suppose we have a thermometer that measures the temperature t within an error $x = t - t_0$ from the true temperature, where x is normally distributed with standard deviation σ.

Suppose we make N independent measurements of the temperature. How are the errors distributed?

$$P(|x_1| < \delta, \ldots, |x_N| < \delta) =$$
$$\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_{x_1=-\delta}^{\delta} \cdots \int_{x_N=-\delta}^{\delta} e^{-(\sum_{i=1}^N x_i^2)/(2\sigma^2)} dx_1 \ldots dx_N$$

This is the *multivariate gaussian* distribution.

t_0 true temp

$x = t - t_0$ t = temp on thermometer.

$$P(|x| < \delta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right) \int_{x=-\delta}^{\delta} e^{-x^2/(2\sigma^2)} dx.$$

Make N independent measurements. $t_1, \ldots, t_N$

$x_i = t_i - t_0.$

$$P(|x_1| < \delta, |x_2| < \delta, \ldots, |x_n| < \delta)$$
$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_{x_1=-\delta}^{\delta} \int_{x_2=-\delta}^{\delta} \cdots \int_{x_N=-\delta}^{\delta} e^{-x_1^2/(2\sigma^2) - x_2^2/(2\sigma^2) \cdots} dx_1 \cdots dx_N$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_{\substack{x_i=-\delta \\ i=1,\ldots,N}}^{\delta} e^{-\|(x_1,\ldots,x_n)\|^2/(2\sigma^2)} dx_1 \cdots dx_n$$

$$P((x_1, \ldots, x_n) \in U \subseteq \mathbb{R}^N) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \int_U e^{-\|x\|^2/2\sigma^2} d\vec{x}$$

Non-independent events

Suppose we draw a pair of real numbers (x, y) from the plane R^2 controlled by the distribution

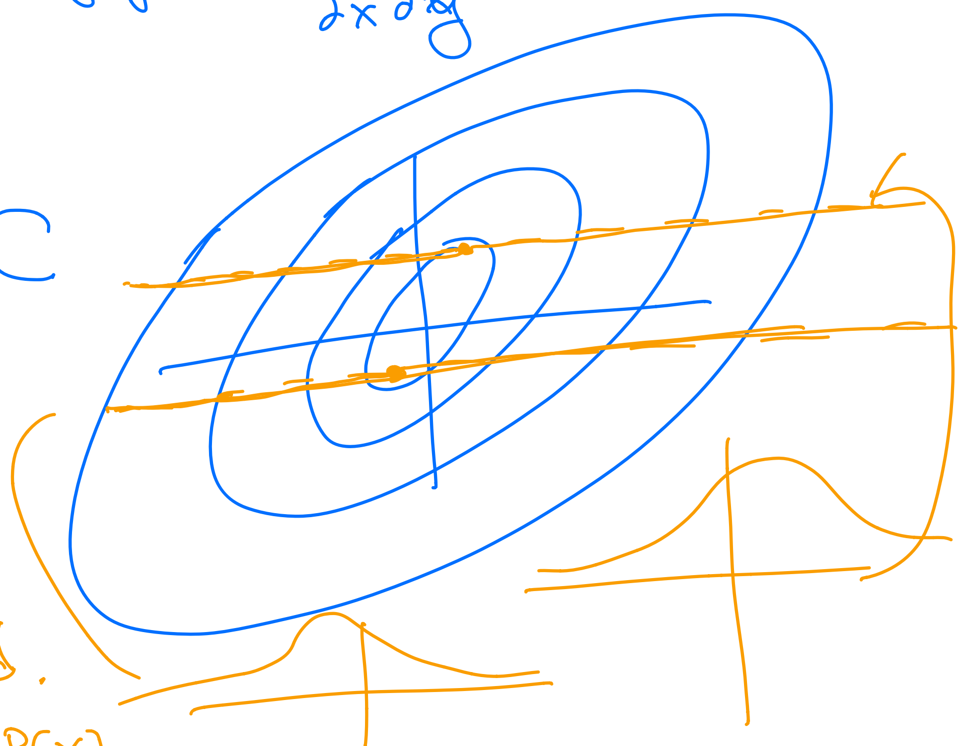
$$P((x,y) \in U) = A \int_U e^{(-x^2-xy-y^2)/(2\sigma^2)} dxdy$$

This density function has a bump at the origin and its level curves are ellipses.

The two coordinates are not independent of each other.

$$P((x,y) \in U) = A \int_U e^{(-x^2-xy-y^2)/2\sigma^2} \, dx\, dy$$

$$x^2 + xy + y^2 = C$$



$P(x|y)$?

x, y are NOT independent

they are correlated.

$P(x|y) = P(x)$

Non-independent events

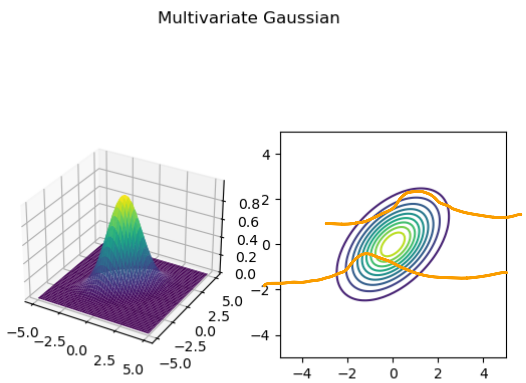


Figure 1: Multivariate Gaussian