2) Probability-Statistics

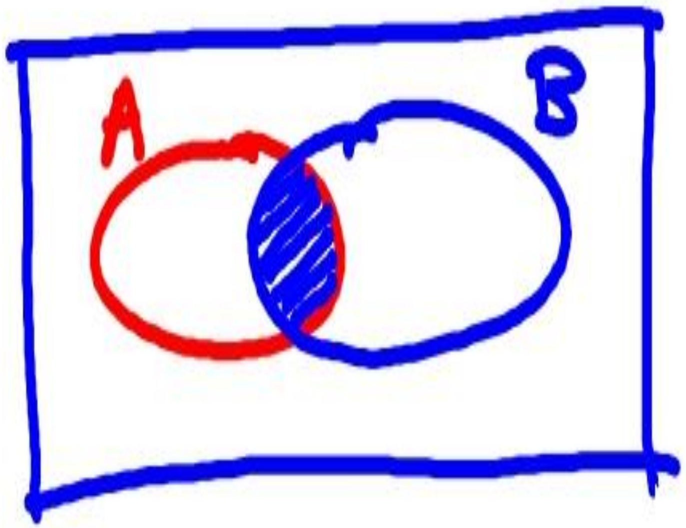
frequency

Probability of A: $P(A) = \lim_{\#\text{trials} \to \infty} \frac{\#\text{ occurences of } A}{\#\text{ trials}}$

$$0 \leq P(A) \leq 1.$$

Conditional Probability

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



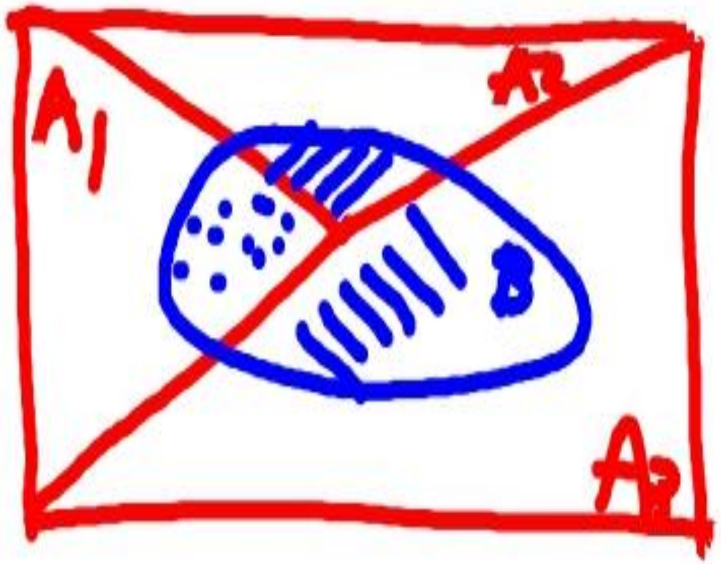
$$P(A|B)P(B) = P(B|A)P(A)$$

$$\Rightarrow \boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)}} \rightarrow \text{Bayes' rule.}$$

Many events A_i

i) Mutually exclusive

ii) Their probabilities add up to 1.



$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \dots + P(A_n \cap B)$$

$$= P(B|A_1)P(A_1) + \dots + P(B|A_n)P(A_n)$$

$$\boxed{P(B) = \sum_i P(B|A_i)P(A_i)} \leftarrow \text{Total probability theorem}$$

Example:

$C=1 \qquad C=2$

$P(A=H \mid C=1) = 0.6 \qquad P(C=1) = 0.7$

$P(A=H \mid C=2) = 0.4 \qquad P(C=2) = 0.3$

a priori probabilities

$$P(C=1 \mid A=H) = \frac{P(A=H \mid C=1)\,P(C=1)}{P(A=H \mid C=1)\,P(C=1) + P(A=H \mid C=2)\,P(C=2)}$$

A posteriori

$$= \frac{0.42}{0.42+0.12} \simeq 0.77$$

Decide that coin 1 was tossed

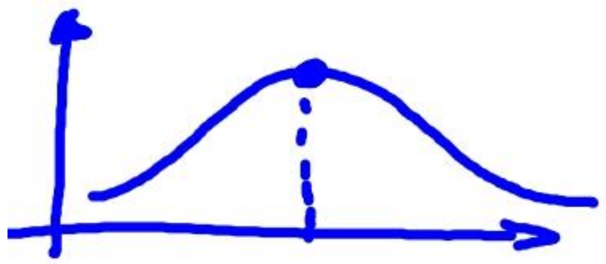
Random variables.

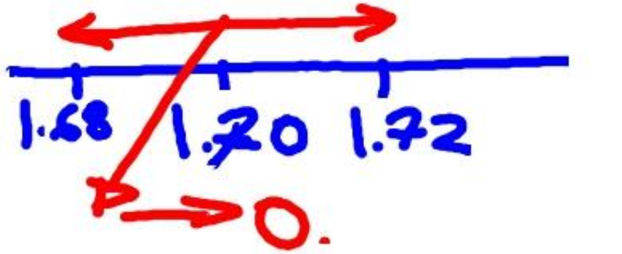
E.g. Z = outcome of throwing a die

1, 2, 3, 4, 5, 6 <u>discrete</u>

X: Height of a person: <u>continuous.</u>

Probability density





A discrete, x, y continuous.

Bayes:

probability density

$$p(x|A)\,\mathbb{P}(A) = \mathbb{P}(A|x)\,p(x)$$

Probability.

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

Total probability.

$$p(x) = \sum_i p(x|A_i)\,\mathbb{P}(A_i)$$

$$\mathbb{P}(A) = \int \mathbb{P}(A|x)\,p(x)\,dx$$

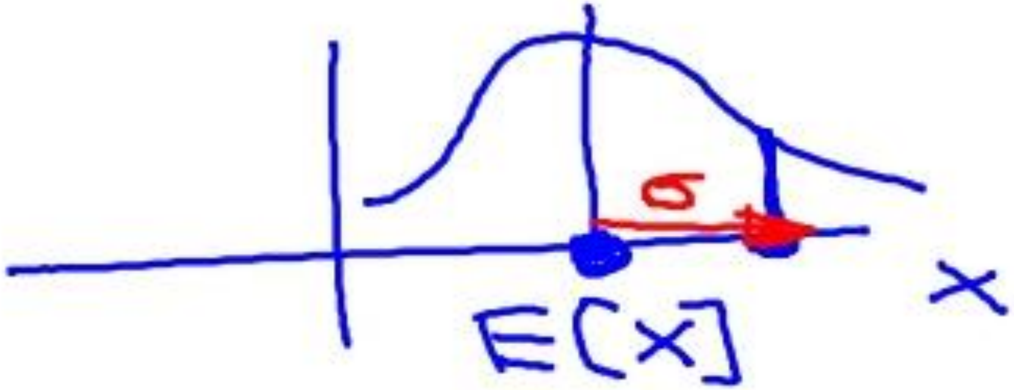
Expectation value:

Discrete: $\mu = E[z] = \sum_i P(z_i) z_i$

Continuous: $\mu = E[x] = \int p(x) \cdot x \cdot dx$.

$a > 0 \quad E[az] = aE[z]$

$E[z_1 + z_2] = E[z_1] + E[z_2]$



Standard deviation:

$\sigma = \sqrt{E[(x-\mu)^2]}$

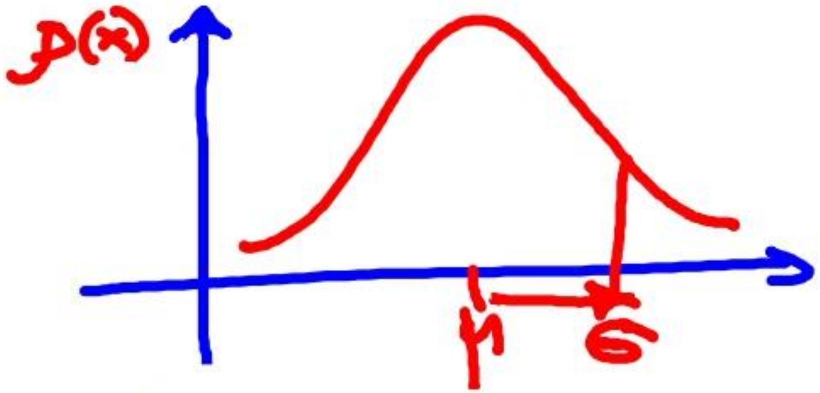
Variance σ^2

Discrete: $\sigma = \sqrt{\sum_i P(z_i)(z_i - \mu)^2}$

Continuous: $\sigma = \sqrt{\int p(x)(x-\mu)^2 dx}$

Gaussian (Normal) distribution

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right]$$



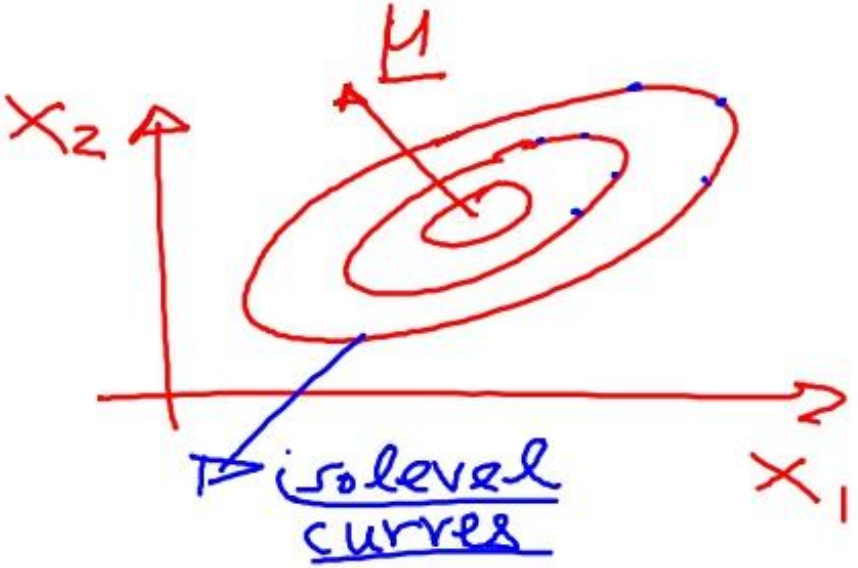
Multidimensional Gaussian:

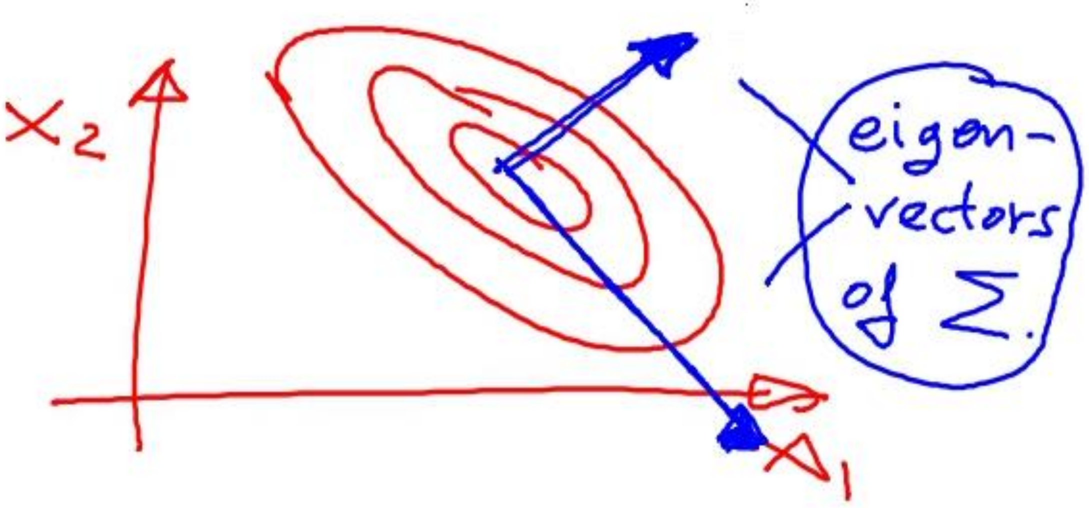
$$p(\underline{x}) = \frac{1}{(2\pi)^{d/2}} \frac{1}{(\det\Sigma)^{1/2}} \exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})\right]$$

dimensions.

Σ : covariance matrix

$$\Sigma = E\left[(\underline{x}-\underline{\mu})(\underline{x}-\underline{\mu})^T\right]$$





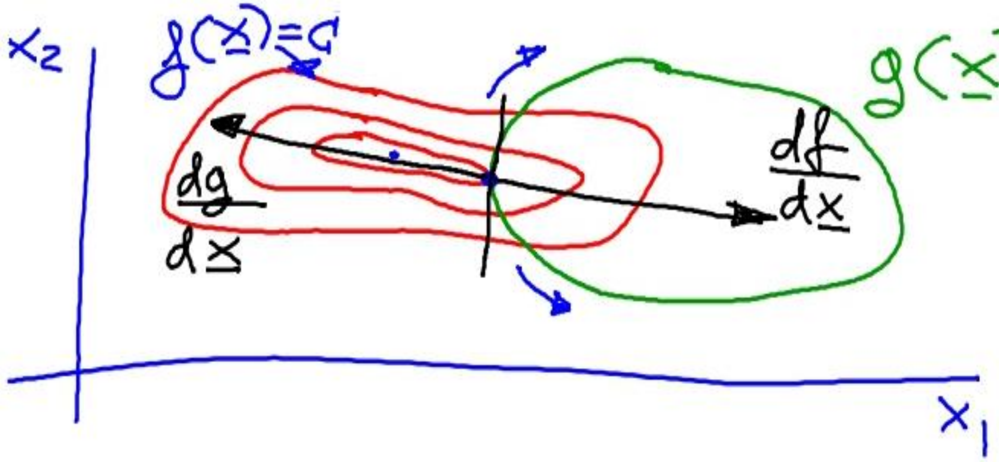
3) Optimization.

Optimize $f(\underline{x})$

$$\frac{df(\underline{x})}{d\underline{x}} = 0.$$

Constrained optimization:

Optimize $f(\underline{x})$ subject to: $g(\underline{x})=0$



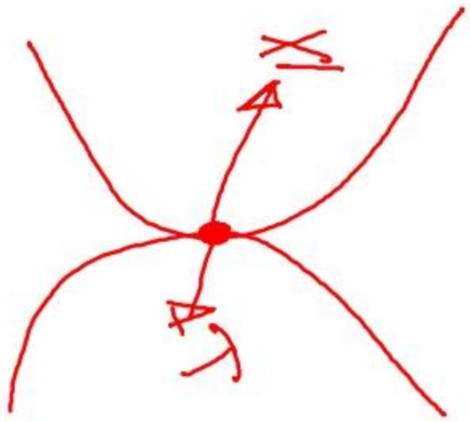
$\frac{df}{d\underline{x}}, \frac{dg}{d\underline{x}}$ collinear:

$$\exists \lambda: \quad \frac{df}{d\underline{x}} + \lambda \frac{dg}{d\underline{x}} = 0 \quad \Rightarrow \frac{d(f+\lambda g)}{d\underline{x}} = 0.$$

$$\mathcal{L}(\underline{x}, \lambda) = f + \lambda g.$$

Lagrangian.

Lagrange multiplier



Optimum point is a saddle point!

Then: $\dfrac{\partial \mathcal{L}}{\partial \underline{x}} = 0, \quad \dfrac{\partial \mathcal{L}}{\partial \lambda} = 0.$

$$\Rightarrow g(\underline{x}) = 0.$$

With many constraints: $g_i(\underline{x})$, $i=1,\dots,v$

$$\mathcal{L} = f + \sum_i \lambda_i g_i$$

Lagrange multipliers

$$\frac{\partial \mathcal{L}}{\partial \underline{x}} = 0 \qquad \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0.$$

$\forall g_i(\underline{x})$.

Example:

Minimize $f = x^2 + y^2$

s.t. $x + y = 1.$

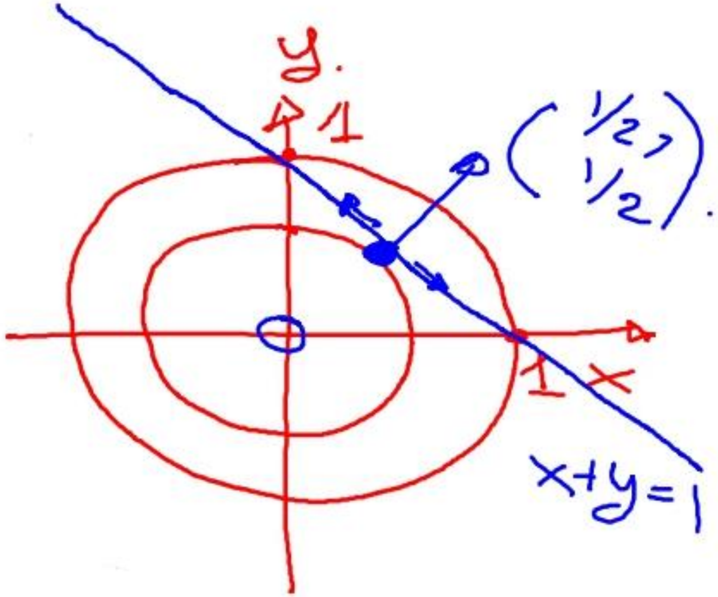
$x + y - 1 = 0.$

$\mathcal{L} = x^2 + y^2 + \lambda(x + y - 1).$

$\frac{\partial \mathcal{L}}{\partial x} = 0 \Rightarrow 2x + \lambda = 0 \Rightarrow x = -\lambda/2$

$\frac{\partial \mathcal{L}}{\partial y} = 0 \Rightarrow 2y + \lambda = 0 \Rightarrow y = -\lambda/2$

$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \Rightarrow x + y - 1 = 0 \Rightarrow x + y = 1$



Therefore: $-\frac{\lambda}{2} - \frac{\lambda}{2} = 1 \Rightarrow \lambda = -1$

$$\Rightarrow \boxed{\begin{array}{l} x = \frac{1}{2} \\ y = \frac{1}{2} \end{array}}$$