Lecture 1 - Introduction to probability theory

Quantum Mechanics is by nature probabalistic:

$$P_i \propto |\langle i | \psi(t) \rangle|^2$$

The normalization of states can be chosen to make this consistent with other definitions of probability.

Frequentist Probability:

Probability that event x occurs is

$$P_x = \lim_{N \to \infty} \frac{\text{Number of times } x \text{ occurs}}{N}$$

Notes: Usually N cannot really go to infinity but it can often be made "large enough".

Only well defined for repeatable experiments.

"What is the probability that $m_H > 100$ GeV?"

does not make sense using frequentist probability.

Bayesian Probability:

$P_x = 1$ corresponds to absolute certainty that x is true.

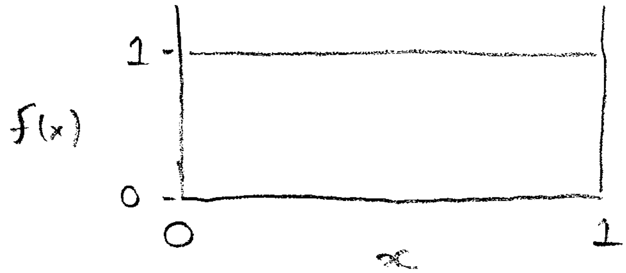
$P_x = 0$ corresponds to absolute certainty that x is not true.

Bayes' theorem describes how our certainty or uncertainty should change when confronted with new information.

Random Variables :

The outcome of drawing from a Probability Distribution.

Example : the uniform distribution.



All values between 0 and 1 are equally probable.

Probability : $P(x) = 0$

Probability density : $\frac{dP(x)}{dx} = f(x) = 1$.

Probability that X is between x and $x+dx$

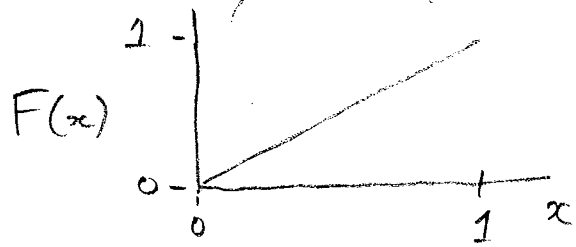
$$P(x < X < x + dx) = \int_x^{x+dx} f(x)\,dx$$

$$= dx$$

Cumulative distribution :

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

" probability that $X < x$.



Other probability density functions:

Gaussian Distribution

$$g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Cumulative distribution:

$$G(x) = \int_{-\infty}^{x} g(x; \mu, \sigma)\, dx$$

$$= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right) \quad \text{(exercise)}$$

Often the result of the "Central Limit Theorem".

Poisson Distribution

$$P(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}$$

Probability of observing n events in an interval (eg of time) when the rate is ν.

When ν is large, $P(n; \nu) \to g(n; \nu, \sqrt{\nu})$

Binomial Distribution.

Probability of "success" is p
Probability of "failure" is $q = 1-p$.

Probability of n successes in N trials is

$$P(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

Describing probability distributions:

Mean, or expected value of x:

$$E(x) = \langle x \rangle = \int x f(x)\, dx$$

$$E(g(x)) = \langle g(x) \rangle = \int g(x) f(x)\, dx$$

Variance of x:

$$\mathrm{Var}(x) = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$$

Standard deviation:

$$\sigma_x = \sqrt{\mathrm{Var}(x)}$$

If x and y are two random variables, their covariance is defined

$$\mathrm{Cov}(x, y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle$$

Correlation coefficient

$$\rho_{xy} = \frac{\mathrm{Cov}(x, y)}{\sigma_x \sigma_y}$$

If $\mathrm{Cov}(x,y) = 0$ then x and y are said to be uncorrelated.

Others: Median, skewness, kurtosis.

Multivariate distributions:

One independent trial yields two (or more) random variables.

Example: Let $V = \mathrm{Cov}(x,y)$.

If V is positive definite then

$$f(x,y) = \frac{1}{2\pi|V|} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T V^{-1} (\vec{x}-\vec{\mu})\right)$$

where $|V| = \det(\mathrm{Cov}(x,y))$

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \vec{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$$

This also generalizes to more than 2 dimensions.

When x and y are uncorrelated,

$$f(x,y) = \frac{1}{(2\pi)\sigma_x\sigma_y} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right)$$

Conditional Probability

Probability of A given the event B :

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

A and B are independent if

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

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

$$f(x,y) = f(x)\,g(y) \quad \text{(PDF's factorize)}$$

Bayes' Theorem :

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

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

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

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad \text{(Bayes' Theorem)}$$

Statistics :

Sample mean $\bar{x} = \frac{\sum_i x_i}{N}$

$\bar{x}$ is a random variable.

$$\langle \bar{x} \rangle = \langle x \rangle$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

(Corrected) sample standard deviation

$$s^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^2$$

s^2 is an unbiased estimate of $Var(x)$.

$$\chi^2 = (\vec{f} - \vec{x})^T V^{-1} (\vec{f} - \vec{x})$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

V is diagonal when x_i are uncorrelated.

Then $V^{-1} = \begin{pmatrix} 1/\sigma_1^2 & & 0 \\ & 1/\sigma_2^2 & \\ 0 & & \ddots \end{pmatrix}$

χ^2 is a random variable. When $\langle x_i \rangle = f_i$,

$$\langle \chi_N^2 \rangle = N$$

$$Var(\chi_N^2) = 2N$$

PDF : $f(x;k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$