Probability cheatsheet

Gavin Band, WHG GMS Programme 2021

1. Probability expresses uncertainty:

$P(X|C) = 0$ => "X is definitely not true"

$P(X|C) = 1$ => "X is definitely true"

Extends basic logic

$0 < P(X|C) < 1$ => we are uncertain about X

2. All probability is conditional:

$\underset{\text{"probability of X"}}{P(X)}$ always means $\underset{\text{"probability of X given C" or "probability of X conditional on C"}}{P(X|C)}$

where **C represents background information** (such as an assumed statistical model). C is often dropped from notation but is always there.

Examples: P(| $\underbrace{\text{die is fair}}_{\text{Model of die}}$) = 1/6 P(transmission | model of infection) etc.

3. Probability is computable:

I. $P(\underset{\text{"not A"}}{!A}|C) = 1 - P(A|C)$ (A either occurs or doesn't.)

II. $P(A \text{ or } B|C) = P(A|C) + P(B|C) - \underset{\text{"A and B given C"}}{P(A,B|C)}$ (One of them must occur, but don't double-count)

III. $\underset{\text{"A and B given C"}}{P(A,B|C)} = \underset{\text{"A given B and C"}}{P(A|B,C)} \cdot \underset{\text{"B given C"}}{P(B|C)}$ (Split up joint probabilities into conditional ones.)

"Independence" happens in when the middle term simplifies: $P(A,B) = P(A)P(B)$

4. Practically useful tools for computation:

(derived directly from the rules in 3)

LOTP: ("Law of total probability")

$$P(A|C) = \sum_X P(A,X|C)$$

Computes probability of A by summing (or integrating) over all the possible values of another variable.

Often seen as $P(A|C) = \sum_X P(A|X,C) \cdot P(X|C)$ by applying rule III

Bayes' rule:

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

Computes P(A given B) in terms of P(B given A)

4. Probability for variables

Often X will instead denote a variable whose value we are uncertain of. It is typical to use lower-case letters like x to denote specific possible values. So we will write:

$P(X = x|C)$ "Probability X takes value x"

For a continuous variable X we must instead think of P as a density function:

$P(X = x|C)$ "*Probability density function* (pdf) evaluated at $X = x$"

and integrate it to get an actual probability:

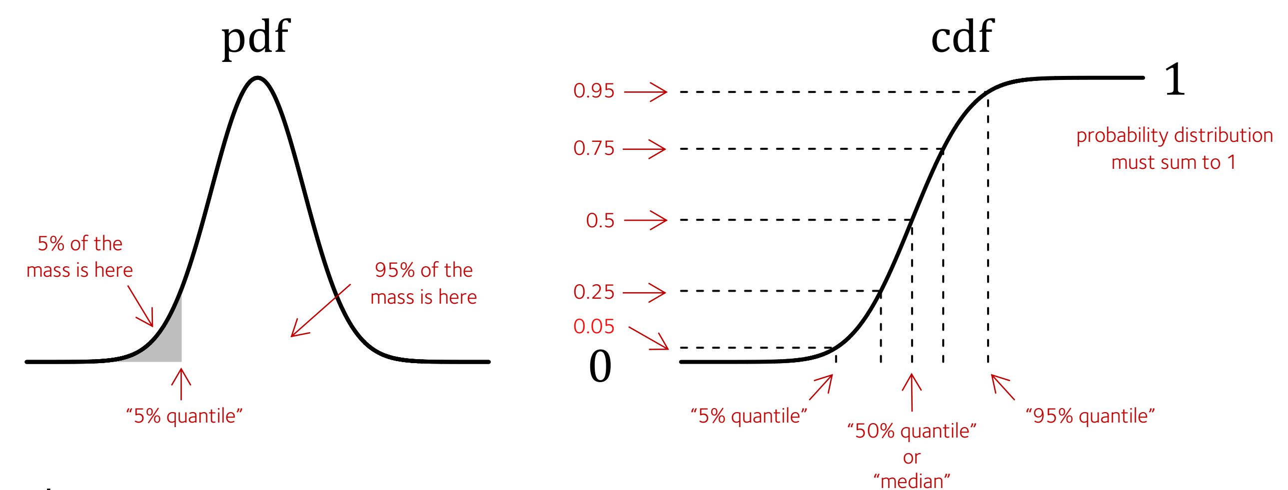
$$P(a \leq X \leq b|C) = \int_a^b P(X = x|C)dx$$

"Probability that X is between a and b"

$$P(X \leq b|C) = \int_{-\infty}^b P(X = x|C)dx$$

"*Cumulative distribution function* (cdf) of X"

Inverse of cdf is the *quantile function*.



5. Distributions

Probabilities are often expressed in terms of a set of well-known distributions

$X \sim N(0,1)$ "X has a standard normal (or Gaussian) distribution"

$X \sim \text{binom}(N,p)$ "X has binomial distribution with N draws and success probability p"

See the separate 'distributions cheatsheet' for more.

6. Statistical summaries of a distribution

"Mean" or *"expected value"* $\mu = E(x) = \sum_x \underset{\text{value}}{x} \cdot \underset{\text{pdf}}{P(X = x)}$ sum (or integrate) over all possible values

"Variance" $\text{var}(x) = E\left((x-\mu)^2\right) = \sum_x (x-\mu)^2 \cdot P(X = x)$ average squared distance to mean value (a measure of uncertainty)

"Entropy" $\text{H}(x) = \sum_x \frac{1}{\log P(X = x)} P(X = x)$ average information provided by x (a measure of uncertainty)