Wikipedia

Thomas Bayes [bɛi:z] (*~1701 –1761) was an English mathematician and Presbytierian minister, known for having formulated a specific case of the theorem that bears his name: Bayes‘ Theorem. Bayes never published what would eventually become his most famous accomplishment.

His work was published posthumously. His ideas gained limited exposure until they were independently rediscovered and further developed by **Laplace** (great French mathematician and astronomer), who first published the modern formulation in his 1812 Théorie analytique des probabilités. Until the second half of the 20th century, the Bayesian interpretation attracted widespread dissent from the mathematics community who generally held frequentist views, rejecting Bayesianism as unscientific. However, it is now widely accepted. This may have been due to the development of computing, which enabled the successful application of Bayesianism to many complex problems.

P(A|B) means:
P(A) <u>if</u> B is given (sure, real) which is synonymous to ...
Given B, A is true with probability P

Example
A states: “Our F_2 individual is Aa”
B states: “Our F_2 individual is with dominant phenotype”
P(A|B) states:
“P is the probability that F2 is Aa **if** the phenotype is dominant”
... is synomymous with saying
“Given its phenotype is dominant, F2 is Aa with probability P”

So: what-now is this probablity … that an F2-plant is heterozygous **if** we know that its phenotype is the dominant phenotype?

A: “Our F_2 individual is Aa” in F_2 P=0.50

B: “Our F_2 individual is with dominant phenotype” in F_2 P=0.75

P(A&B): Probability A & B simultaneously true in F_2 P=0.50

Conditional probability tells us:

$P(A|B) = P(A\&B) / P(B)$

P(Aa **if** dom. phenotype) = (½) / (¾) = ⅔

Message:

“⅔ of F_2-plants with dominant phenotype are heterozygous”

synonymous

“Given dominant phenotype, genotype is Aa with P = ⅔”

A: “Our F_2 individual is Aa”
B: “Our F_2 individual is with dominant phenotype”
P(A&B): Probability that in F2, A and B are true simultaneously

Conditional probability tells us:

$P(A|B) = P(A\&B) / P(B)$. Thomas Bayes says ☺: $P(A|B) \cdot P(B) = P(A\&B)$

$P(B|A) = P(A\&B) / P(A)$. Thomas Bayes says ☺: $P(B|A) \cdot P(A) = P(A\&B)$

Therefore (Thomas Bayes’ ingenious idea):

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

So he isolated P(A|B):

$P(A|B) = P(B|A) \cdot P(A) / P(B)$

That’s it, all else is just bells and whistles …  

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

Without Bayes you would say:

„⅔ of those F2 plants with dominant phenotype are heterozygous".

Calculation: P(A|B) = (½) / (¾) = ⅔

You use the **JOINT probability** „dominant and heterozygous simultaneously" which is **½**.

You calculate P(A|B) = **P(A&B)** / P(B)

You calculate P(A|B) = (½) / (¾) = ⅔

Bayes can do this without **P(A&B)** = (½) … only with

P(B|A) = Probability phenotype dominant if Aa, this is 100%

P(A) = Probability that genotype in F2 is Aa, this is 50%

P(B) = probability that phenotype is dom. in F2, this is 75%

P(A|B) = P(B|A) · P(A) / P(B) = 1.0 · 0.5 / 0.75 = $0,\overline{6}$

BAYESIAN POPULATION GENETIC DATA ANALYSIS

Kent E. Holsinger
Department of Ecology & Evolutionary Biology, U-3043
University of Connecticut
Storrs, CT 06269-3043

You can find very many extremely instructive texts on population genetics and quantitative genetics if you search for **Kent Holsinger**

Motivation for Bayesian inference

Suppose a man is known to have transmitted A_1 to his offspring at a locus having only two alleles, A_1 and A_2. What genotype is he likely to have had?

The likelihood of the data, given each of the possible paternal genotypes is:*

$$\begin{aligned} \mathrm{P}(x = A_1 | \phi = A_1A_1) &= 1 \\ \mathrm{P}(x = A_1 | \phi = A_1A_2) &= 1/2 \\ \mathrm{P}(x = A_1 | \phi = A_2A_2) &= 0 \end{aligned}$$

Using likelihood as our criterion, we'd always conclude that the paternal genotype is A_1A_1.

*x is a random variable referring to the allele that the father transmits to his offspring. ϕ is a random variable referring to the genotype of the father.

Motivation for Bayesian inference

Suppose a man is known to have transmitted A_1 to his offspring at a locus having only two alleles, A_1 and A_2. What genotype is he likely to have had?

The likelihood of the data, given each of the poss[ible] genotypes is:*

$$P(x = A_1 | \phi = A_1A_1) = 1$$
$$P(x = A_1 | \phi = A_1A_2) = 1/2$$
$$P(x = A_1 | \phi = A_2A_2) = 0$$

Given that he is A1A1, he transmits A1 with P=1.0

Given that he is A1A2, he transmits A1 with P=0.5

I would say – from this reasoning – that this man‘s genotype is probably A1A1. The probability that he is A1A1 is 2/3 and for A1A2 it is 1/3, isn‘t it?

Motivation for Bayesian inference

But suppose we knew that the frequency of the A_1 allele in the population this man belonged to was only 1%.

- Then the frequency of the A_1A_1 genotype is only 0.0001.
- Nearly all of the A_1 alleles in the population are carried in heterozygotes, so it doesn't seem likely that the father would have been A_1A_1.
- Most of the offspring carrying A_1 would have inherited it from a heterozygous father.
- Likelihood would be misleading!

98%

Bayes' Theorem

Likelihood of the data, x, given a particular model and set of parameters, ϕ

$$L(\phi|x) \propto \mathrm{P}(x|\phi)$$

The likelihood of the data is proportional to the probability of getting the data, given the unknown parameter. So if we are interested in the probability that the parameter (ϕ, the father's genotype) takes a particular value given the data (x, the off-spring's genotype), we can just apply Bayes' Theorem

$$\mathrm{P}(\phi|x) = \frac{\mathrm{P}(x|\phi)\mathrm{P}(\phi)}{\mathrm{P}(x)}$$

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

Motivation for Bayesian inference

But how do we make that intuitive conclusion precise? We want to calculate the probability that the father had a particular genotype *given* that he transmitted allele A_1 to his offspring.

Any time you hear the word "given" you should think conditional probability, and when you think "conditional probability," you should think "Bayes Theorem."

Given that he is A1A1, he transmits A1 with P=1.0

We want to calculate these probabilities:*

$$P(\phi = A_1A_1 | x = A_1)$$
$$P(\phi = A_1A_2 | x = A_1)$$
$$P(\phi = A_2A_2 | x = A_1)$$

Given that he transmitted A1, he is A1A1 with P=?

*Remember that x is a random variable referring to the allele that the father transmits to his offspring. ϕ is a random variable referring to the genotype of the father.

$$P(A|B) = P(B|A) \cdot P(A) / P(B)$$

Bayes' Theorem

$$P(\phi|x) = \frac{P(x|\phi)P(\phi)}{P(x)}$$

- $P(\phi)$: **prior probability**, the probability you would assign to the parameter ϕ taking a particular value *before* you look at any data.*

- $P(\phi|x)$: **posterior probability**, the probability you assign to the parameter ϕ given the observed data, x.

*So where does this come from? Good question. We'll get to that in a moment.

Motivation for Bayesian inference

Returning to our earlier example:

$$P(A|B) = P(B|A) \cdot P(A) \;/\; P(B)$$

$$\mathrm{P}(\phi = A_1A_1 | x = A_1) = \frac{\mathrm{P}(x = A_1 | \phi = A_1A_1)\mathrm{P}(\phi = A_1A_1)}{\mathrm{P}(x = A_1)}$$

Probability that he is “A1A1”, given he transmitted “A1” to offspring = Probability that he transmits “A1” given he is “A1A1” times Probability that he is “A1A1” (without condition) divided by Probability to transmit “A1” (without condition).

$$P(A|B) = P(B|A) \cdot P(A) / P(B)$$

$$P(\phi = A_1A_1|x = A_1) = \frac{P(x = A_1|\phi = A_1A_1)P(\phi = A_1A_1)}{P(x = A_1)}$$

Probability that he is “A1A1”, given he transmitted “A1” to offspring

... we need

Probability that he transmits “A1” given he is “A1A1”
This is for sure, P=1.0

A1A2 0.5

Probability that he is “A1A1” (without condition)
This is (HWE) p^2, here this is 0.01^2 which is 1 in 10.000

A1A2 0.0198

Probability to transmit “A1” (without condition)
this is p=0.01. The probability that a randomly chosen allele in the population is A1, without any condition, is just its frequency p(A1)=0.01

$$P(\phi = A_1A_1|x = A_1) = \frac{P(x = A_1|\phi = A_1A_1)P(\phi = A_1A_1)}{P(x = A_1)}$$

$$P(A|B) = P(B|A) \cdot P(A) / P(B)$$

So the posterior probabilities are

$$P(\phi = A_1A_1|x = A_1) = \frac{(1)(p_1^2)}{p_1} = p_1 = 0.01$$

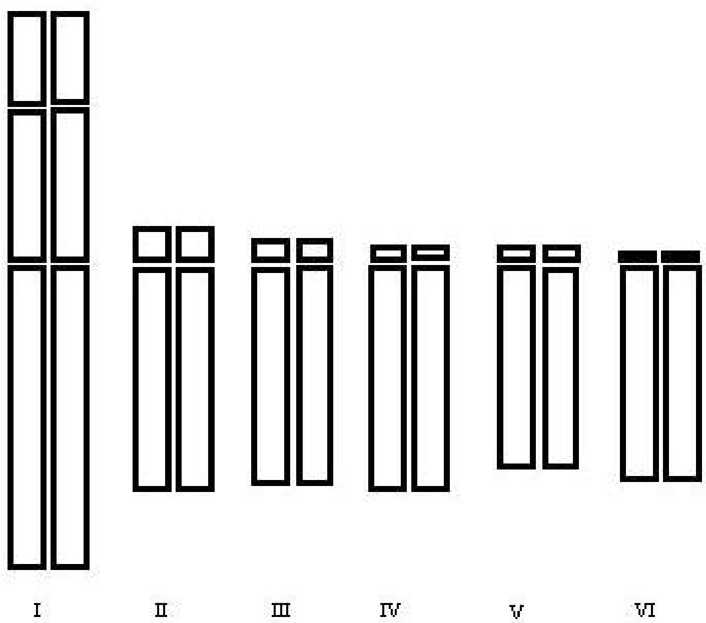
Suppose a man is known to have transmitted A_1 to his offspring at a locus having only two alleles, A_1 and A_2. What genotype is he likely to have had?

He was either A_1A_1 or A_1A_2,.

But: The likelihood that he was A_1A_1 is only **1%** whereas

the likelihood that he was A_1A_2 is 99%

I am quite sure that Bayesian methods are e.g. used to construct linkage maps. In this case, for grouping marker loci to linkage groups, it would be reasonable to a priori „tell“ the software the number of expected groups (such as x=6 in case of *Vicia faba*). It is not really necessary to again and again - when constructing a map - behave as if we did not know the actual number of chromosomes. This is - I think - a convincing example of using **prior** information.



http://yudkowsky.net/rational/bayes; FUN FACTS

Q. What is the Bayesian Conspiracy?
A. The Bayesian Conspiracy is a multinational, interdisciplinary, and shadowy group of scientists that controls publication, grants, tenure, and the illicit traffic in grad students. The best way to be accepted into the Bayesian Conspiracy is to join the Campus Crusade for Bayes in high school or college, and gradually work your way up to the inner circles. It is rumored that at the upper levels of the Bayesian Conspiracy exist nine silent figures known only as the Bayes Council.

Q. Where do priors *originally* come from?
A. Never ask that question.

Q. Uh huh. Then where do scientists get their priors?
A. There's a **small, cluttered antique shop in a back alley of San Francisco's Chinatown**. *Don't ask about the **bronze rat**.*

Q. Are there any limits to the power of Bayes' Theorem?
A. According to legend, one who fully grasped Bayes' Theorem would gain the ability to create and physically enter an alternate universe using only off-the-shelf equipment and a short computer program.

... cluttered antique shop in a back alley of San Francisco's Chinatown (alias diagon alley; '*Winkelgasse*').

A tourist wanders into a back alley antique shop in San Francisco's Chinatown. Picking through the objects on display he discovers a detailed, life-sized bronze sculpture of a rat. The sculpture is so interesting and unique that he picks it up and asks the shop owner what it costs.

"Twelve dollars for the rat, sir," says the shop owner, "and a thousand dollars more for the story behind it."

"You can keep the story, old man," he replies, "but I'll take the rat."

The transaction complete, the tourist leaves the store with the bronze rat under his arm. As he crosses the street in front of the store, two live rats emerge from a sewer drain and fall into step behind him. Nervously looking over his shoulder, he begins to walk faster, but every time he passes another sewer drain, more rats come out and follow him. By the time he's walked two blocks, at least a hundred rats are at his heels, and people begin to point and shout.

He walks even faster, and soon breaks into a trot as multitudes of rats swarm from sewers, basements, vacant lots, and abandoned cars. Rats by the thousands are at his heels, and as he sees the waterfront at the bottom of the hill, he panics and starts to run full tilt.

No matter how fast he runs, the rats keep up, squealing hideously, now not just thousands but millions, so that by the time he comes rushing up to the water's edge, a trail of rats twelve city blocks long is behind him. Making a mighty leap, he jumps up onto a light post, grasping it with one arm while he hurls the bronze rat into San Francisco Bay with the other, as far as he can heave it.

Pulling his legs up and clinging to the light post, he watches in amazement as the seething tide of rats surges over the breakwater into the sea, where they drown.

Shaken and mumbling, he makes his way back to the antique shop.

"Ah, so you've come back for the rest of the story," says the owner.

"Ah, so you've come back for the rest of the story," says the owner.

"No," says the tourist, "I was wondering if you have a bronze lawyer."