Bayesian Inference

Pere Puig

Department de Matemàtiques

Universitat Autònoma de Barcelona

ppuig@mat.uab.cat

Overview

1- Bayes' theorem and its consequences.

2- The basics of Bayesian Statistics: prior distributions.

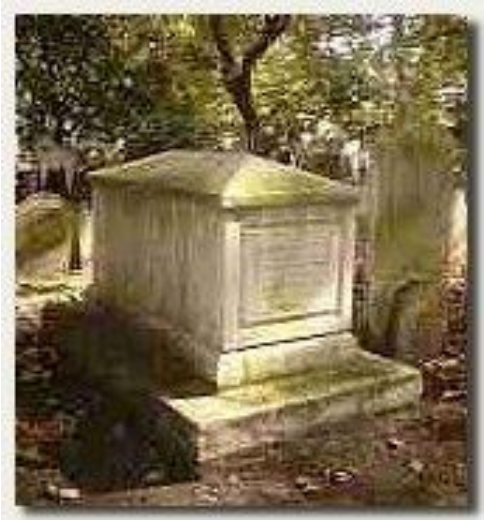
3- Bayesian inference: the posterior distribution.

4- Bayesian hypothesis testing.

Thomas Bayes 1702-1761

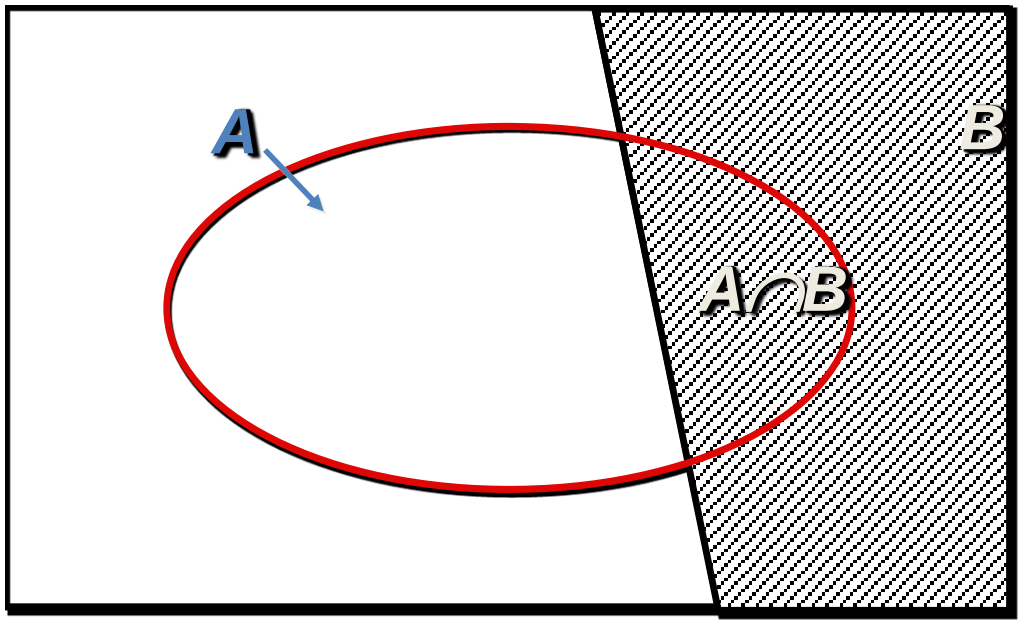
Two years after his death, "An Essay Towards Solving a Problem in the Doctrine of Chances" by English clergyman and mathematician Thomas Bayes,was published in the Philosophical Transactions of the Royal Society 53 (1763) 370-418. This paper was sent to the Royal Society by Bayes's friend Richard Price. Bayes's paper enunciated Bayes's Theorem for calculating "inverse probabilities".

In recognition of Thomas Bayes's important work in probability. The vault was restored in 1969 with contributions received from statisticians throughout the world.



Conditioned probability

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



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

Conditioned probability incorporates the *prior* information (B) to calculate the probability of A.

Enjoy with these exercises!

- Mr. Jones has two children. What is the probability that both children are girls? (no prior information)
- Mr. Jones has two children. At least one of them is a girl. What is the probability that both children are girls?
-Mr. Jones has two children. The older child is a girl. What is the probability that both children are girls?
-Read the history of the professor of statistics carrying a bomb in his baggage:
http://www.onlinemathlearning.com/math-jokes-statistics.html

Enjoy with these exercises!

- We are to play a version of Russian Roulette, the revolver is a standard six shooter but I will put two bullets in the gun in consecutive chambers. I spin the chambers, put the gun to my head pull the trigger and survive. I hand you the gun and give you a choice... You may put the gun straight to your head and pull the trigger, or you may re-spin the gun before you do the same.

  What is the better choice in order to survive?

1- Bayes' theorem

Using the expressions of conditioned probability,

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

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

and combined them,

$$P(A)P(B \mid A) = P(B)P(A \mid B)$$

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)}$$

Frequently, the probability of the denominator is calculated using the law of total probability, obtaining this expression:

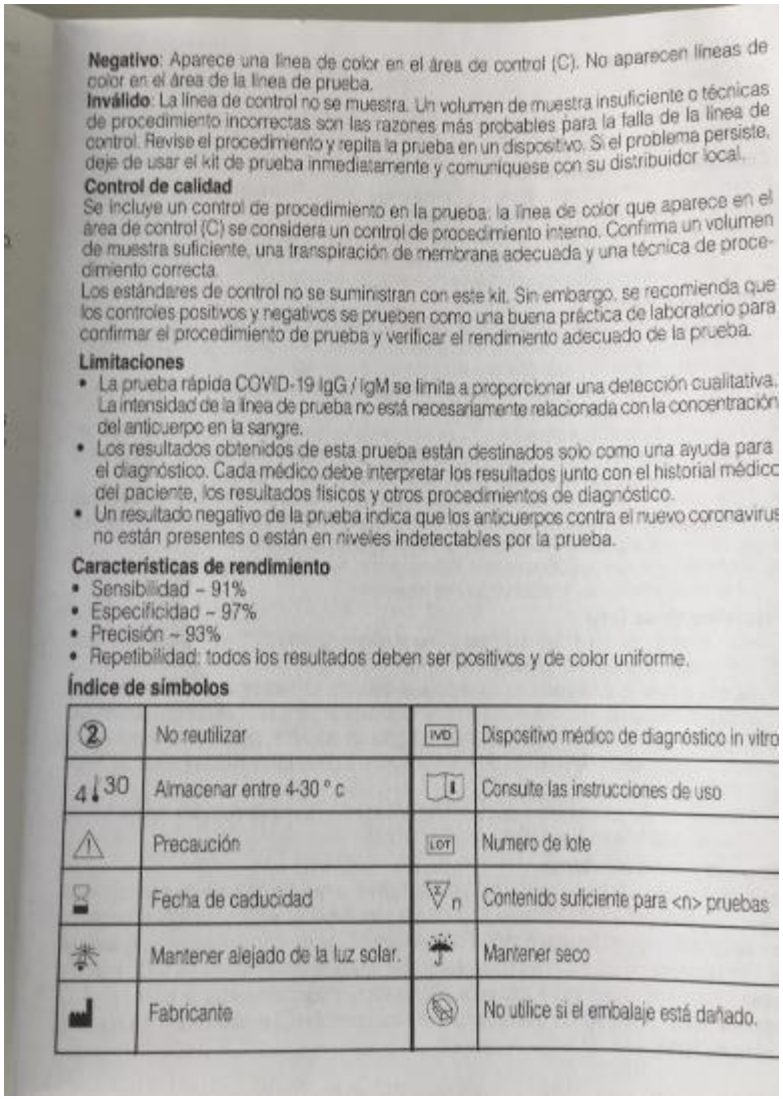
$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)}$$

Example

A diagnostic test helps to calculate the probability for an individual to have a disease.

- We know initially the value of P(sick) because this is the prevalence of the disease.
  - Prevalence: The proportion of individuals in a population having a disease.
- We use a dignostic or clinical test to aid in the diagnosis or detection of disease. The result of the diagnostic test can be “positive” (+) or “negative” (-). Each test is evaluated, and the following parameters of quality are known:
  - P(+ | sick)= Sensibility = the ability of the test to identify positive results
  - P(- | healthy)= Specificity = the ability of the test to identify negative results
- Using Bayes’ theorem we can calculate the predictive values of the diagnostic test: P(sick | +) and P(healthy | -).

Rapid test for SARS-CoV-2 diagnosis



Negativo: Aparece una línea de color en el área de control (C). No aparecen líneas de color en el área de la línea de prueba.
Inválido: La línea de control no se muestra. Un volumen de muestra insuficiente o técnicas de procedimiento incorrectas son las razones más probables para la falla de la línea de control. Revise el procedimiento y repita la prueba en un dispositivo. Si el problema persiste, deje de usar el kit de prueba inmediatamente y comuníquese con su distribuidor local.

Control de calidad
Se incluye un control de procedimiento en la prueba: la línea de color que aparece en el área de control (C) se considera un control de procedimiento interno. Confirma un volumen de muestra suficiente, una transpiración de membrana adecuada y una técnica de procedimiento correcta.
Los estándares de control no se suministran con este kit. Sin embargo, se recomienda que los controles positivos y negativos se prueben como una buena práctica de laboratorio para confirmar el procedimiento de prueba y verificar el rendimiento adecuado de la prueba.

Limitaciones
- La prueba rápida COVID-19 IgG / IgM se limita a proporcionar una detección cualitativa. La intensidad de la línea de prueba no está necesariamente relacionada con la concentración del anticuerpo en la sangre.
- Los resultados obtenidos de esta prueba están destinados solo como una ayuda para el diagnóstico. Cada médico debe interpretar los resultados junto con el historial médico del paciente, los resultados físicos y otros procedimientos de diagnóstico.
- Un resultado negativo de la prueba indica que los anticuerpos contra el nuevo coronavirus no están presentes o están en niveles indetectables por la prueba.

Características de rendimiento
- Sensibilidad ~ 91%
- Especificidad ~ 97%
- Precisión ~ 93%
- Repetibilidad: todos los resultados deben ser positivos y de color uniforme.

Índice de símbolos

②	No reutilizar	IVD	Dispositivo médico de diagnóstico in vitro
4 30	Almacenar entre 4-30 ° c		Consulte las instrucciones de uso
⚠	Precaución	LOT	Numero de lote
	Fecha de caducidad	n	Contenido suficiente para <n> pruebas
	Mantener alejado de la luz solar.		Mantener seco
	Fabricante		No utilice si el embalaje está dañado.

Exercise: Compute P(sick |+) and P(healthy |-) using the values of the estimated prevalence of SARS-CoV-2 in Barcelona (7% approx.)

https://www.mscbs.gob.es/ciudadanos/ene-covid/docs/ESTUDIO_ENE-COVID19_INFORME_FINAL.pdf

The Bayesian thinking

- Before (prior) the rapid covid test, the chance of a patient to have covid is just 7% (the population prevalence).

- After (posterior) the new data (the result of the diagnostic test has been +), the chance is 0.695

Bayesian methods learn from data. The role of the experiment is to add to our knowledge and so to update what we can say about the parameters.

Bayesian statistics versus non-bayesian

Statistics estimates unknown parameters (like the mean or the variance of a population). Parameters represent things that are unknown. They are some properties of the population from which the data arise. Questions of interest are expressed as questions on such parameters: confidence intervals, hypothesis testing etc.

Classical (frequentist) statistics considers parameters as specific to the problem, so that they are not subject to random variability. Hence parameters are just unknown constants, they are not random, and it is not possible to make probabilistic statements about parameters. (Remember the meaning of a CI)

Conversely, Bayesian statistics considers parameters as random and hence it is allowed to make probabilistic statements about them. The distribution of the parameters is called *prior distribution*, expressing what is known (or believed to be true) before seeing the data.

2-The prior distribution

Approaches to choose a prior distribution:

- Informative prior distribution. We have to use the knowledge about the problem, perhaps based on other data. It is necessary to construct a distribution reflecting all beliefs about the unknown parameters.

- Noninformative prior distribution. This is implemented by assigning equal probability to all possible values of the parameters (uniform distributions).

The prior distribution: an example

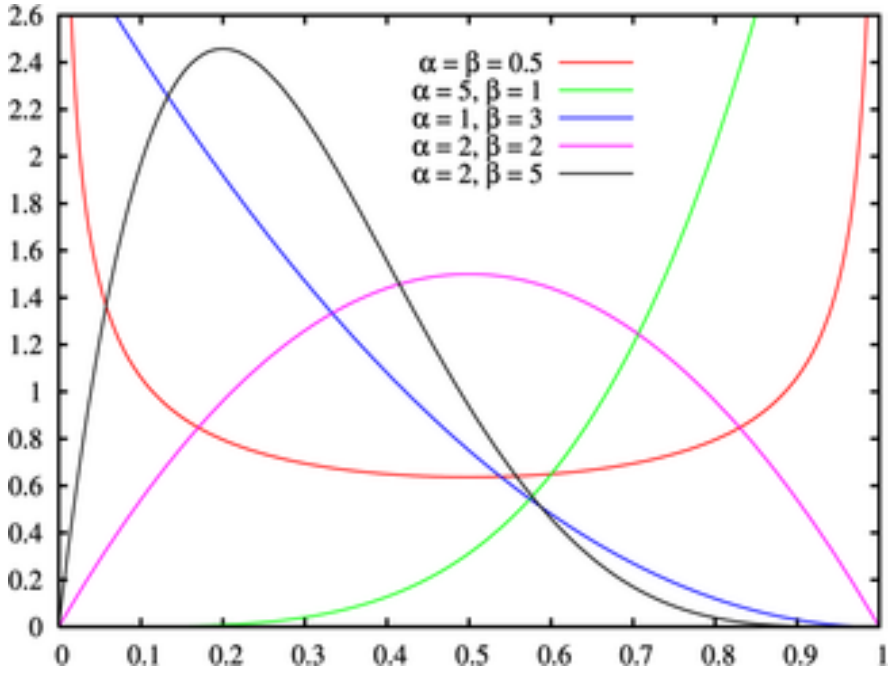
(from *Basic Bayesian Methods*, Mark E. Glickman and David A. van Dyk)

We are interested in the probability θ, that a randomly selected woman, aged 60–69 years with a family history of cancer, had a positive breast cancer screening. Because the parameter θ belongs to [0,1], we have to define a prior distribution over this interval. What we know *a priori* about θ?

If we do not trust in the information provided by the American Cancer Society, a natural noninformative prior distribution is the uniform over [0,1]. But an expert tell us, "Anyway I'm sure that the value of θ can not be greater than 5%". Then we could define a kind of noninformative (??) prior distribution using the uniform over [0,0.05].

However, according to the American Cancer Society, roughly 3.6% of women aged 60–69 years develop invasive breast cancer, so that we can construct an informative prior distribution for θ reflecting this information. In this case, a flexible choice could be the Beta distribution.

Beta densities (https://en.wikipedia.org/wiki/Beta_distribution)



$$f(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,du}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}$$

$$= \frac{1}{\mathrm{B}(\alpha,\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}$$

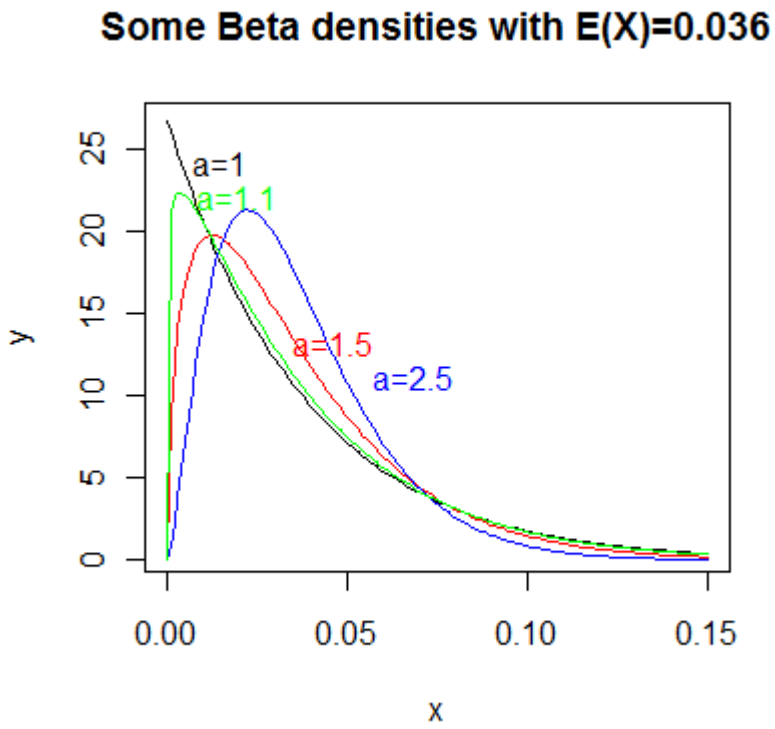
The expectation and variance are,

$$\mathrm{E}(X) = \frac{\alpha}{\alpha+\beta}$$

$$\mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

We can take $\beta = 26.7778\alpha$ and then $E(X) = 0.036$. Using values $\alpha > 1$ the density has a mode.

```
x=seq(0,0.15,0.001)
y=dbeta(x,1,1*26.7778)
y2=dbeta(x,1.5,1.5*26.7778)
y3=dbeta(x,2.5,2.5*26.7778)
y4=dbeta(x,1.1,1.1*26.7778)
plot(x,y,type="l")
lines(x,y2,col="red")
lines(x,y3,col="blue")
lines(x,y4,col="green")
legend(0.002,26,"a=1",bty="n")
legend(0.052,13,"a=2.5",text.col="blue",bty="n")
legend(0.03,15,"a=1.5",text.col="red",bty="n")
legend(0.003,24,"a=1.1",bty="n",text.col="green")
title(main = Some Beta densities with E(X)=0.036)
```



3- Bayesian Inference: the posterior distribution

Once the data has been observed, the likelihood function is constructed. This is the joint probability function of the data, but viewed as a function of the parameters. Assuming that the data values, $X= (x_1, \ldots, x_n)$ are obtained independently, the likelihood function is given by,

$$L(X|\theta) = p(x_1, \ldots, x_n|\theta) = \prod_{i=1}^{n} p(x_i|\theta)$$

where $p(x|\theta)$ is the density or probability function for each observation.

To obtain the *posterior distribution (or density)*, $f(\theta|X)$, the probability distribution of the parameters once the data have been observed, we apply Bayes'theorem.

The posterior distribution

$$f(\theta \mid X) = \frac{\pi(\theta) L(X \mid \theta)}{\int_{-\infty}^{\infty} \pi(\theta) L(X \mid \theta) d\theta}$$

Here $\pi(\theta)$ is the prior density. Note that the denominator is just a normalizing constant not depending on the parameter.

It is straightforward to obtain the posterior distribution: Simply multiply the prior distribution by the likelihood, and then determine the constant that forces the expression to integrate to 1.

From the posterior density we can estimate θ by means of the expectation (or median, or mode), and we can calculate confidence intervals, called *credibility intervals*, using the percentiles.

An example (continued)

(from *Basic Bayesian Methods*, Mark E. Glickman and David A. van Dyk)

Suppose, for the breast cancer screening study, 14 of the 300 women had positive tests. Thus 14 women have $x_i = 1$, and the remaining 286 have $x_i = 0$. The likelihood is therefore given by, $L(X|\theta) = \theta^{14}(1-\theta)^{286}$

Taking the Beta prior distribution with parameters $\alpha=2.5$ and $\beta=66.9445$, the following posterior density is obtained: $f(\theta|X) = c\,\theta^{15.5}(1-\theta)^{351.9445}$
Thus, the posterior distribution is Beta(16.5 , 352.9445).
Then we can obtain point estimates of the parameter in the following way:

1- Expectation	$\theta^*_1 = 0.0447$	$\alpha/(\alpha+\beta)$
2- Mode	$\theta^*_2 = 0.0422$	$(\alpha-1)/(\alpha+\beta-2)$
3- Median	$\theta^*_3 = 0.0438$	qbeta(0.5,16.5,352.9445)
4- Classical (non-Bayesian)	$\theta^*_4 = 0.0467$	

An example (continued)

The likelihood is therefore given by, $L(X|\theta)= \theta^{14}(1-\theta)^{286}$

Taking the uniform[0,1] prior noninformative distribution, the following posterior density is obtained: $f(\theta|X) = c\,\theta^{14}(1-\theta)^{286}$
Thus, the posterior distribution is Beta(15 , 287).
Then we can obtain point estimates of the parameter in the following way:

1- Expectation	$\theta^*_1 = 0.0497$	$\alpha/(\alpha+\beta)$
2- Mode	$\theta^*_2 = 0.0467$	$(\alpha-1)/(\alpha+\beta-2)$
3- Median	$\theta^*_3 = 0.0487$	qbeta(0.5,15,287)
4- Classical (non-Bayesian)	$\theta^*_4 = 0.0467$	

"Confidence intervals" can be calculated using the percentiles. For example, a 95% central posterior interval involves computing the 2.5% and 97.5% percentiles of the posterior distribution. Taking the uniform[0,1] noninformative distribution, the corresponding percentiles are,

```
qbeta(c(0.025,0.975),15,287)
[1] 0.02815644      0.07680793
```

Thus, a 95% credibility interval for θ is [0.0281 , 0.0768] .
Taking the Beta prior distribution with parameters α=2.5 and β=66.9445, the percentiles are now,

```
qbeta(c(0.025,0.975),16.5,352.9445)
[1] 0.02606071      0.06791458
```

Now, a 95% credibility interval for θ is [0.0261 , 0.0679]

The classical 95% confidence interval is,

$0.0467 \pm 0.0238 = [0.0229 , 0.0705]$

Interpretation: There is a 95% chance that the proportion θ lies between 2.3% and 7.1%.

Correct? NO!

However, in the Bayesian context, parameters are random and when we compute a Bayesian interval for θ it means exactly the interpretation usually wrongly given to a confidence interval.

Summary: the steps to perform Bayesian inference.

1. Formulate a probability model for the data.

2. Decide on a prior distribution, which quantifies the uncertainty in the values of the unknown model parameters before the data are observed.

3. Observe the data, and construct the likelihood function based on the data and the probability model formulated in step 1. The likelihood is then combined with the prior distribution from step 2 to determine the posterior distribution, which quantifies the uncertainty in the values of the unknown model parameters after the data are observed.

4. Summarize important features of the posterior distribution, or calculate quantities of interest based on the posterior distribution. These quantities constitute statistical outputs, such as point estimates and intervals.

Exercise: Make Bayesian inference for the breast cancer screening study.

-Use the opinion of the expert, “Anyway I’m sure that the value of θ can not be greater than 5%”, and use as a prior distribution the uniform over [0 , 0.05].

-Calculate the posterior density and find point estimates of parameter θ.

-Calculate a 95% centered credibility interval for θ.

4- Bayesian hypothesis testing

In classical (frequentist) statistics the p-value is just the probability of rejecting the null hypothesis (H_0) given the null hypothesis is really true. It can be written in this way:

$$\text{p-value} = P(\text{ Data } | \ H_0 \text{ is true})$$

However, the p-value is frequently misinterpreted as $P(\ H_0$ is true $|$ Data) !!

But $P(\ H_0$ is true $|$ Data) is very interesting and meaningful. Bayesian inference allows to compute this probability and, in fact, this is the Bayesian procedure for hypothesis testing.

We want to test $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_1$

The distribution of the data given the parameter: $X|\theta \sim L(X|\theta)$

Prior distributions:

-Prior probabilities of the hypothesis, $P(H_0)$ and $P(H_1)$. An uninformative option is $P(H_0)=P(H_1)=0.5$.
-Prior densities $\pi_0(\theta)$ and $\pi_1(\theta)$ defined on Θ_0 and Θ_1. If Θ_i is a point then the density $\pi_i(\theta)$ is a point mass.

Posterior probabilities: (an application of Bayes Theorem)

$$P(H_0|X) = \frac{P(H_0)\int_{\Theta_0} L(X|\theta)\pi_0(\theta)d\theta}{P(H_0)\int_{\Theta_0} L(X|\theta)\pi_0(\theta)d\theta + P(H_1)\int_{\Theta_1} L(X|\theta)\pi_1(\theta)d\theta} = 1 - P(H_1|X)$$

Example: Remember the breast cancer screening study, 14 of the 300 women had positive tests. We want to test,

H_0: $\theta = \theta_0$ (θ_0 =0.036)
H_1: $\theta \neq \theta_0$

We assume that $P(H_0)=P(H_1)=0.5$, and $\pi_1(\theta)$ to be uniform.

$$P(H_0|X) = \frac{\int_{\Theta_0} L(X|\theta)\pi_0(\theta)d\theta}{\int_{\Theta_0} L(X|\theta)\pi_0(\theta)d\theta + \int_{\Theta_1} L(X|\theta)\pi_1(\theta)d\theta} = \frac{L(X|\theta_0)}{L(X|\theta_0) + \int_{\theta\neq\theta_0} L(X|\theta)d\theta}$$

Note that due to continuity, $\int_{\theta\neq\theta_0} L(X|\theta)d\theta = \int_0^1 L(X|\theta)d\theta$

Then, $\int_0^1 L(X|\theta)d\theta = \int_0^1 \theta^s(1-\theta)^{n-s}d\theta = \beta(s+1, n-s+1)$,
where $s = \sum_{i=1}^n x_i$

Consequently,

theta=0.036;x=14;n=300
1/(1+theta^(-x)*(1-theta)^(x-n)*beta(x+1,n-x+1))
$P(H_0|X)$ = 0.9541

Exercise:

For the example of the breast cancer screening study,
Perform the following test,

H_0: $\theta = 0.05$
H_1: $\theta < 0.05$

-Calculate $P(H_0|X)$ and $P(H_1|X)$.

The Bayes factor

We are interested in testing H_0 against H_1. In order to make a decision we calculate the proportion,

$$\frac{Pr(H_0|X)}{Pr(H_1|X)} = \frac{Pr(H_0|X)}{1 - Pr(H_0|X)}$$

This proportion is known as *Bayes factor*. This is just the posterior odds in favor of H_0 over H_1. If we have that $P(H_0)=P(H_1)=0.5$ and the hypothesis are simple this is just a likelihood ratio.

Note that for the breast cancer screening study, $P(H_0|X) = 0.9541$ and the Bayes factor is, 21.786. What is the conclusion?

Table 1. Jeffreys' scale of evidence.

$B_{12} \in$	Evidence against M_2
(1,3.2)	Not worth more than a bare mention
(3.2,10)	Substantial
(10,100)	Strong
(100,∞)	Decisive.

Table 2. Scale of evidence proposed by Kass and Raftery (1995).

$B_{12} \in$	Evidence against M_2
(1,3)	Not worth more than a bare mention
(3,20)	Positive
(20,150)	Strong
(150,∞)	Very strong.

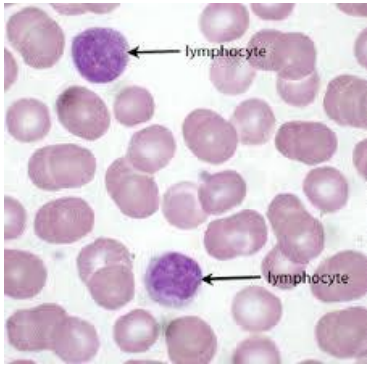
In these tables M_2 is our H_1. Although these rules are given in terms of evidence against H_1, the same rule can be used to interpret the evidence against H_0 by inverting the value of the Bayes factor. Here the Bayes factor is indicated as B_{12}.

Example 1

A clinical test is performed by means of a blood analysis, where the number of lymphocytes are counted in a certain number of squares. An individual is considered healthy if the mean is H_0: $2 \leq \mu \leq 4$ and infected if H_1: $\mu > 4$. The result of a clinical test has given a sample mean of 4.6 with a sample of 100 squares. Assuming that the counts are Poisson distributed, what is the conclusion?

Note that for Poisson observations the likelihood function is,

$$L(\lambda|x_1, ..., x_n) \propto P(x_1, ..., x_n|\lambda)$$
$$\propto \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}$$



Assuming that $P(H_0)=P(H_1)=0.5$, the Bayes factor is,

$$\frac{Pr(H_0|X)}{Pr(H_1|X)} = \frac{\int_2^4 L(\mu|X)\pi_0(\mu)\,d\mu}{\int_4^\infty L(\mu|X)\pi_1(\mu)\,d\mu}$$

Assuming $\pi_0(\mu)= \pi_1(\mu)$ ~Gamma(α,1), where the hyper-parameter α is estimated by the sample mean (4.6), the Bayes factor is,

$$\frac{Pr(H_0|X)}{Pr(H_1|X)} = \frac{\int_2^4 \mu^{463.6}\exp(-101\mu)\,d\mu}{\int_4^\infty \mu^{463.6}\exp(-101\mu)\,d\mu}$$

Numerical calculations can be problematic. However, dividing numerator and denominator by an appropriate constant this is just, (F(4)-F(2))/(1-F(4)) where F is the distribution function of a Gamma(464.6 ,101).
It gives a Bayes factor of,

```
> (pgamma(4,464.6,101)-pgamma(2,464.6,101))/(1-pgamma(4,464.6,101))
[1] 0.001707005
```

It gives an evidence against H_0 of 1/0.001707 = 585.8231.

What is the required sample size in order to obtain an evidence against H_0 ≥ 100, when the sample mean is ≥5?

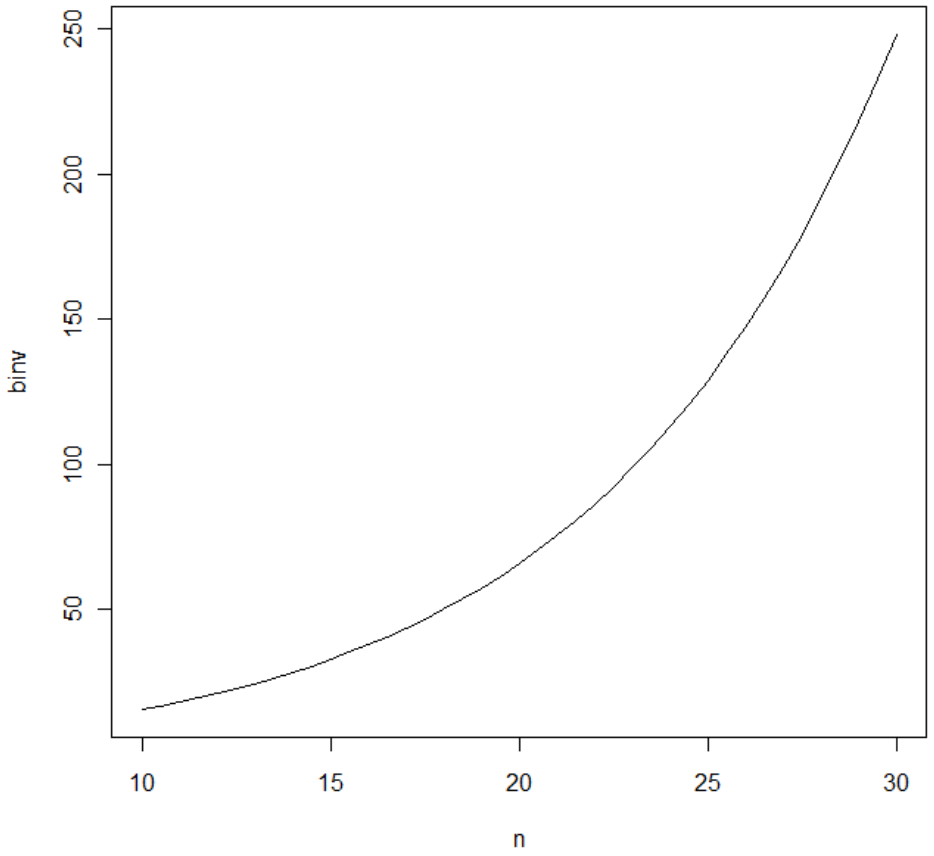
Now the Bayes factor, as a function of the sample size n, is the following:

$$\frac{Pr(H_0|X)}{Pr(H_1|X)} = \frac{\int_2^4 \mu^{5n+4} \exp(-n-1)\, d\mu}{\int_4^\infty \mu^{5n+4} \exp(-n-1)\, d\mu}$$

Again, dividing numerator and denominator by an appropriate constant this is just, (F(4)-F(2))/(1-F(4)) where F is the distribution function of a Gamma(5(n+1),n+1).
We plot the inverse of the Bayes factor against n in order to explore what happens:

```
n=seq(10,30,0.5)
binv=(1-pgamma(4,5*(n+1),n+1))/(pgamma(4,5*(n+1),n+1)-pgamma(2,
5*(n+1),n+1))
plot(n,binv,type="l")
```

The conclusion is that a sample size about 25 can be enough.



Example 2

A subject is asked to guess the colour (red or black) of each of 20 randomly selected playing cards. The subject correctly guesses 15 colours. Does this level of success indicate that she has extra-sensory perception (ESP)?

Under the null hypothesis H_0 that he/she does not have ESP, the probability of correctly identifying 15 out of 20 colours is $L(X|\ H_0) = 0.0148$ (check this!).

If he/she does have ESP (the alternative hypothesis H_1), then we suppose that she has a probability $\theta > 0.5$ of correctly guessing each card. Under the assumption that all values of θ from 0.5 to 1 are equally likely (uniform prior distribution), the probability that the medium correctly identifies 15, given that he/she does have ESP, is

```
integrand= function(t) {dbinom(15,20,t)*2}
integrate(integrand,0.5, 1)$value
```

0.0940

The Bayes factor against the null hypothesis is 0.0940/0.0148 = 6.351. This is clearly not strong evidence against H_0, and would not be enough to persuade most people to believe in ESP.
(Perform a classical test and compare the results.)

What happens if the prior distribution is changed?

Exercise: Prove that, based on the data of 15 correct answers out of 20, then the posterior odds against the null hypothesis will always be lower or equal than 13.736.

Working example: quantification of radiation dose.

Natasja Weitsz / Getty

Alexander Litvinenko in the intensive care unit of University College London Hospital on November 20, 2006, three days before his death from radiation poisoning.

The mysterious circumstances surrounding the death of Litvinenko, apparently from radiation poisoning, spawned an international crisis. Britain demanded that Russia extradite a Russian citizen allegedly connected to the case. When it refused, Britain expelled four Russian diplomats from London, in reprisals reminiscent of the Cold War.

https://www.youtube.com/watch?v=KpReRcw3mf0

Litvinenko case

- **London**

1 Emirates Stadium: Traces of radioactive material found at Arsenal's ground, where ex-KGB man Vyacheslav Sokolenko attended a match during his stay at the Millennium Hotel
2 University College Hospital, where Litvinenko died on 23 November. Mario Scaramella, who met Litvinenko at sushi bar, also treated there
3 Central London: Litvinenko met associates at Millennium Hotel and at Itsu sushi bar on 1 Nov
4 Heathrow Airport: Radioactive material found in three aircraft frequently used in flights between London and Moscow, since given all-clear

Moscow

- Traces of radioactive material found on a Finnish plane and at British Embassy, visited by Andrei Lugovoi, one of those who met Litvinenko in London
- Another contact Litvinenko met at the Millennium Hotel - Dmitry Kovtun - treated for radioactive poisoning

Hamburg

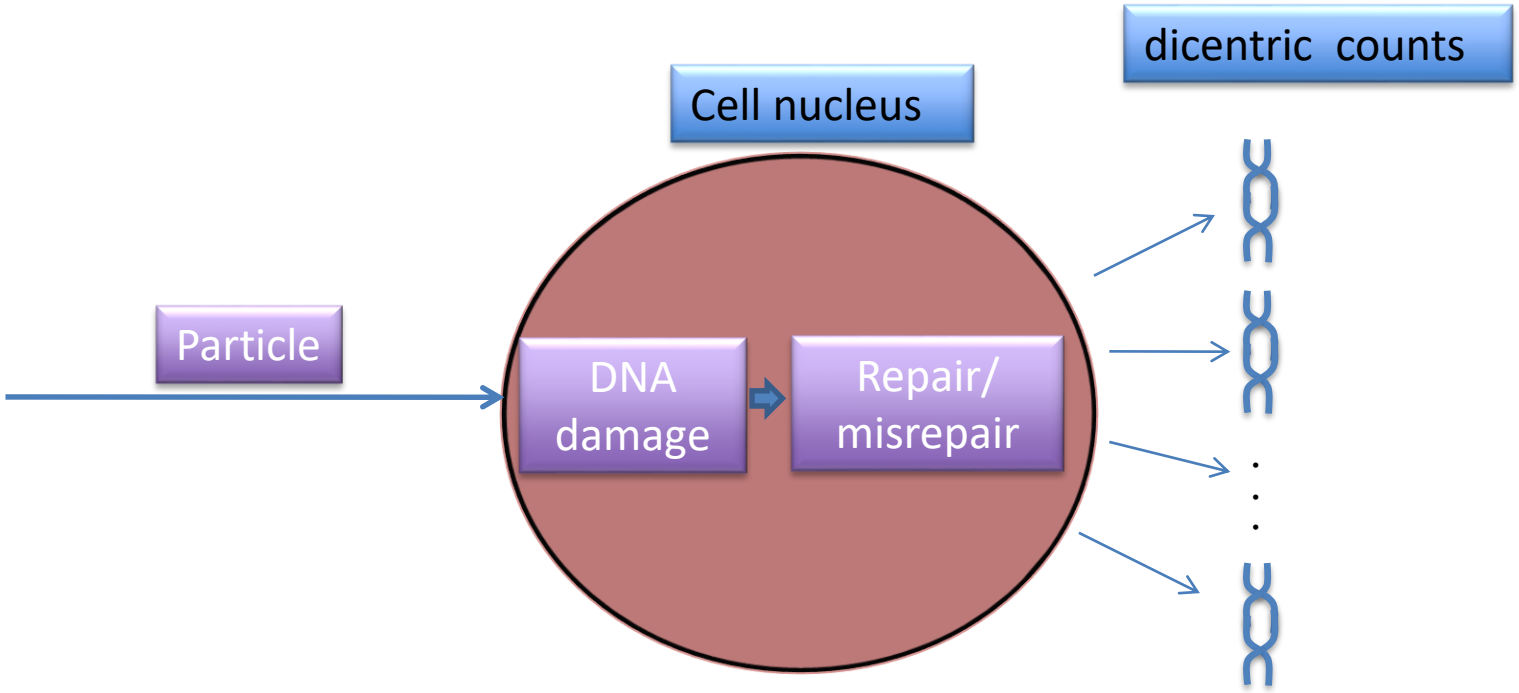
- More radioactive material found in flat of Kovtun's ex-wife, where he had stayed before flying to London, in a car he used and at his mother-in-law's home
- She, her new partner and her two children are contaminated

More than 150 persons were examined to detect possible radioactive contamination.

Quantification of the radiation dose received

This is essential for predicting the derived health consequences, for both early effects (vomiting, skin Injuries, etc.) and late effects like the development of cancer.

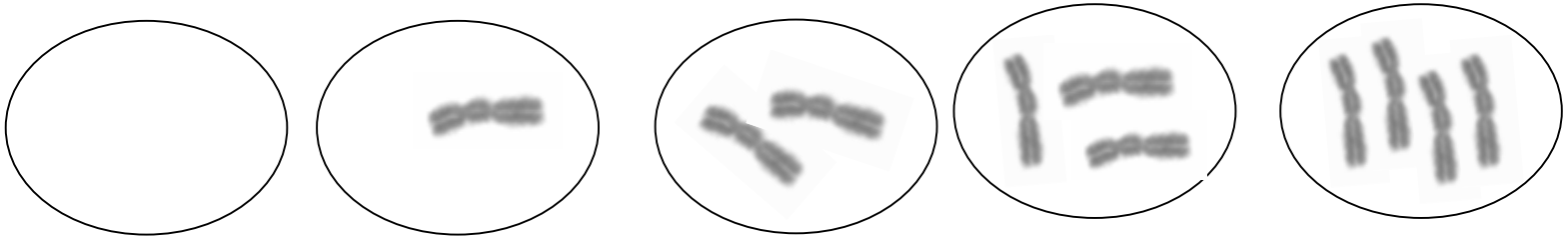
The most widely used method is the analysis of the frequency of dicentrics observed in metaphases from in peripheral blood lymphocytes.



Dicentric formation process

Uniformly stained metaphase spread with dicentric chromosomes (arrows).

For dicentrics, irradiation with X or gamma rays produces a distribution of damage which is very well represented by the Poisson distribution



$$\frac{\lambda^0 e^{-\lambda}}{0!} \qquad \frac{\lambda^1 e^{-\lambda}}{1!} \qquad \frac{\lambda^2 e^{-\lambda}}{2!} \qquad \frac{\lambda^3 e^{-\lambda}}{3!} \qquad \frac{\lambda^4 e^{-\lambda}}{4!}$$

Parameter λ is closely related with the dose of radiation.

Exercise: We have observed 100 lymphocytes obtaining the following frequency distribution of dicentrics:

No. of dicentrics	0	1	2	3
No. of lymphocytes (frequency)	42	28	22	8

Assuming that the data are Poisson distributed, calculate the posterior distribution of the parameter λ using a Gamma distribution as a prior.

- Use the sample mean and variance of the data for choosing an informative Gamma prior. Calculate the posterior distribution of λ providing also bayesian point estimates and a credibility interval.
- If $\lambda > 1.2$ the radiation accident can be considered as "serious". Test de hypothesis H_0: $\lambda \leq 1.2$ and H_1:$\lambda > 1.2$,
  a) when we don't have any initial preference: $P(H_0)=P(H_1)=0.5$
  b) when we have additional Medical evidences that the accident has been serious: $P(H_0)=1/3$, $P(H_1)=2/3$