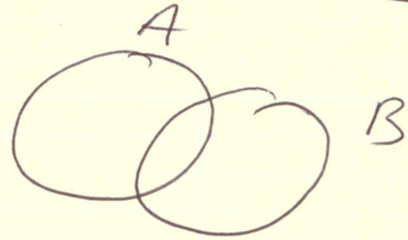
Bayesean Statistics



Two events A and B

$$p(A \cap B) = p(A) \cdot p(B|A)$$

$$p(A \cap B) = p(B) \cdot p(A|B)$$

prob of B happening given A

$$\Rightarrow \boxed{p(B|A) = \frac{p(A|B)\, p(B)}{p(A)}}$$ Bayes Theorem

A measurement of a physics quantity is reported as (say) $v = 10 \pm 1$ m/sec

1σ, 68.3%

What does it mean?

If statement made by a Bayesean:

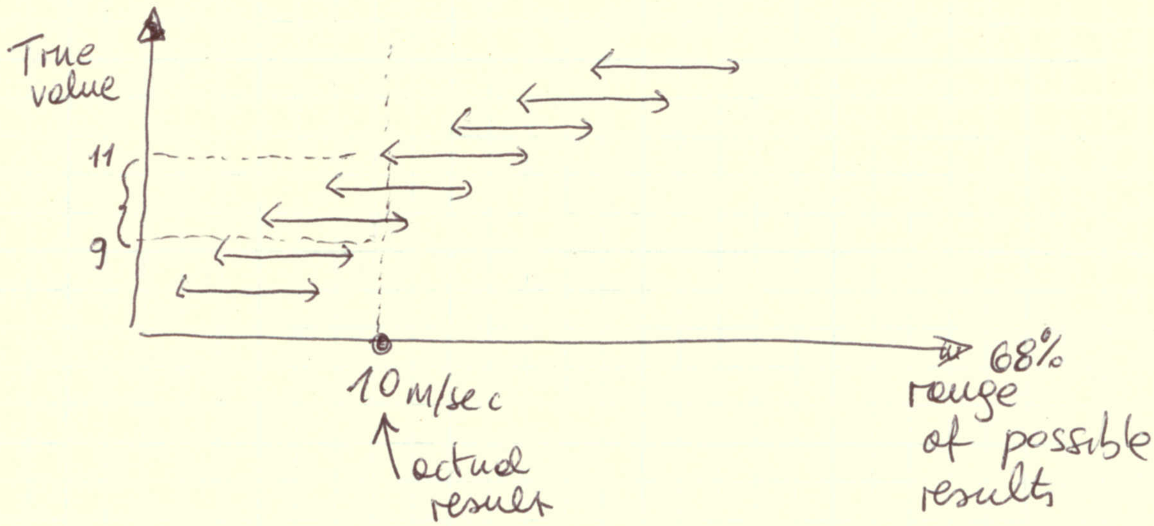
"The probability of the <u>true</u> value of v is 68% that it is between 9 and 11 m/sec"

If the statement made by Frequentist:

"True values between 9 and 11 m/sec are such that they would result in a range of result such that the measured value of 10 is within the 68% range of probable values"

"The probability of the true value is meaningless or in any case we dont talk about it. There is only <u>one</u> true value"

STAT 2



Poisson process. Mean (true) = μ
Observed N

~~P(N|μ)=~~

$$p(N|\mu) = \frac{\mu^N e^{-\mu}}{N!}$$

We have observed N. We want to make a statement about μ (clearly $\mu \sim N$)
Let's be Bayesean.

We want $p(\mu|N)$

Probability of μ given that we have seen N

Bayes theorem

$$P(\mu \mid N) = \frac{P(N \mid \mu)\ P(\mu)}{P(N)}$$

$$P(\mu \mid N) = \frac{\mu^N e^{-\mu}}{N!}\ \frac{P(\mu)}{P(N)}$$

What is $P(N)$? Who knows – It's some number
But it doesnt matter – I have now
<u>seen</u> N – So $P(N)$ is fixed – In fact
$N!$ is also fixed

So $P(\mu \mid N) = \text{Const}\ \mu^N e^{-\mu}\ P(\mu)$

What is $P(\mu)$?
It is my "prior" ~~po in~~ probability
for μ – Reflects my ~~prior~~ knowledge
"prior" to the measurement –
~~What Tees~~ What should I take?
It is ~~an ext~~ a choice that <u>will</u>
<u>affect</u> my answer for $P(\mu \mid N)$

Common choice – I am ignorant –
~~P(μ) can take~~ μ can be anything

STAT 4

Take uniform between 0 and ∞

$$\boxed{P(\mu|N) = \text{Cons}\ \mu^N e^{-\mu}\, d\mu} \quad \left(\begin{array}{c}\text{FLAT}\\ \text{PRIOR}\end{array}\right)$$

μ is real continous variable
N is discrete integer.

Makes kind of sense, but is subjective
Even claiming total ignorance is not unique. I could say that my ignorance could be "parametrized" by saying that μ^2 could take any value uniform between $0 \to \infty$

$$P(\mu) \sim d\mu^2 = 2\mu\, d\mu$$

Would lead to $P(\mu|N) = \text{Const}\ \mu^{N+1} e^{-\mu} d\mu$

This is why many people hate Bayesean statistics.

Nevertheless, let's continue.

Imagine that my counts are a combination of "signal" S and "background" B

$$N = S + B$$

Suppose I know the <u>average</u> B perfectly. But it will be subject to stat. fluctuation. I want some statement about S

$$p(\mu | N) = \text{Const}\ \mu^N e^{-\mu} d\mu$$

$\mu = S + B$

I know this

this is what I care about ← I want it's prop distribution

$$P(S+B | N) = \text{Const}\ (S+B)^N e^{-(S+B)} d(S+B)$$

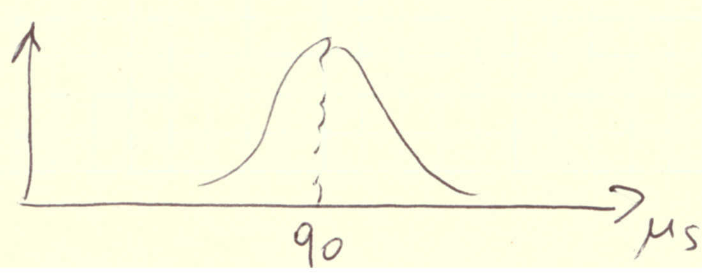
"Relabel" $S \to \mu_s$

$$P(\mu_s | N) = \text{Const}\ (\mu_s + B)^N e^{-(\mu_s+B)} d\mu_s$$

And ~~w~~ actually write it as

$$p(\mu_s | N, B) = \text{Const}\ (\mu_s+B)^N e^{-(\mu_s+B)} d\mu_s$$

Imagine $B=10$ $N=100$, intuitively $\mu_s = 90$ and the $p(\mu_s | N, B)$ will look like this

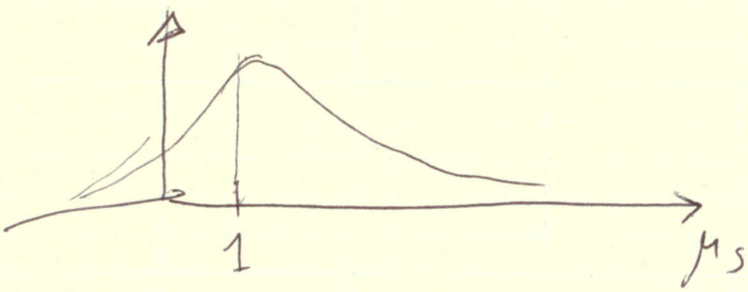


In fact since we know that Poisson → Gaussian for large N we will have something like

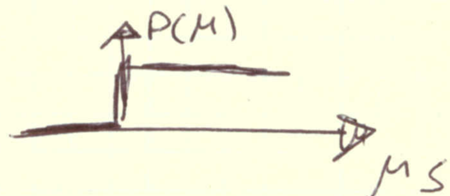
$$\mu_s + B = 100 \pm \sqrt{10} \qquad B = 10$$

$$\mu_s = 90 \pm \sqrt{10} \checkmark$$

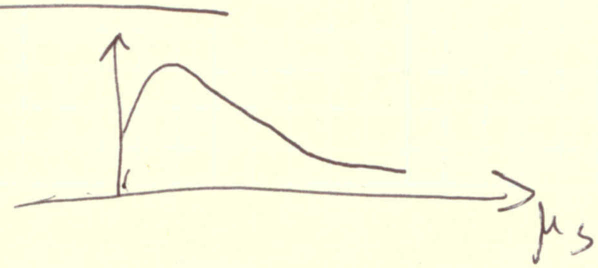
What happens when N close enough to B that it is consistent with just being a fluctuation of B (eg $B=2$, $N=3$



The curve can extend with some area at $\mu_s < 0$ - BUT remember we had multiplied by a prior



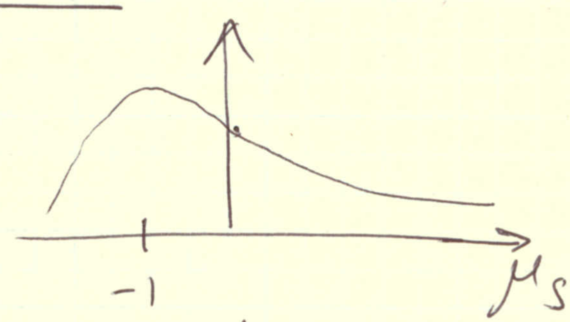
So effectively the curve gets truncated at 0



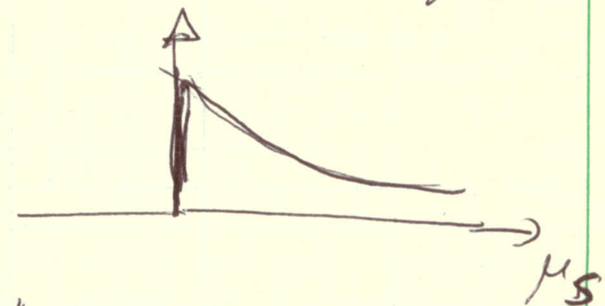
What if eg $B=3$ $N=2$?

The curve would actually ~~peak~~ peak at $\mu_s = -1$ <u>UNPHYSICAL</u>

still truncate it



becomes



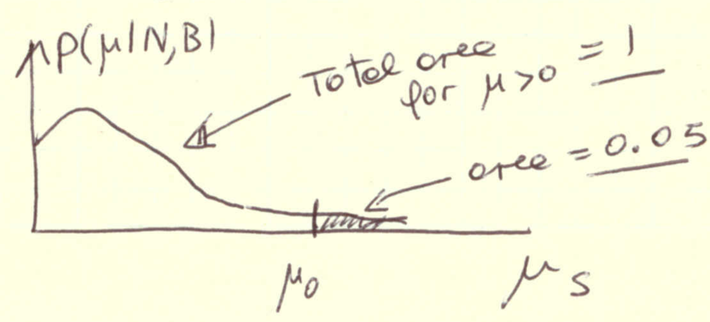
Want then to set "limit" on μ.

Something like: "Given N, B, the probab. that μ_s is greater than μ_0 is 95%

↑ or 90% or whatever you like

Why would you want to do that? Because even if you have not seen any evidence for signal, if signal was "large enough" you would have seen it!

Prescription



STAT 8

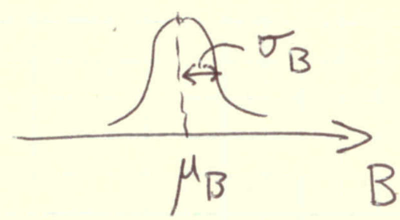
What if you do not know B exactly?

We had

$$P(\mu_s|N,B) = C\,(\mu_s+B)^N e^{-\mu_s+B}\, d\mu_s$$

Assign a PDF to B -

Say a Gaussian



$$P(B)\,dB = G(B|\mu_B,\sigma_B)$$

$$P(\mu_s|N,B) = C\,(\mu_s+B)^N e^{-(\mu_s+B)}\, e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}}\, d\mu_s\, dB$$

B = "nuisance parameter"

The Baysean prescription is to integrate it away (<u>marginalize</u>)

$$P(\mu_s) = C\int_0^\infty dB\,(\mu_s+B)^N e^{-(\mu_s+B)}\, e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}}\, d\mu_s$$

Then you do the same thing again -

You have done this integral in Homework 4