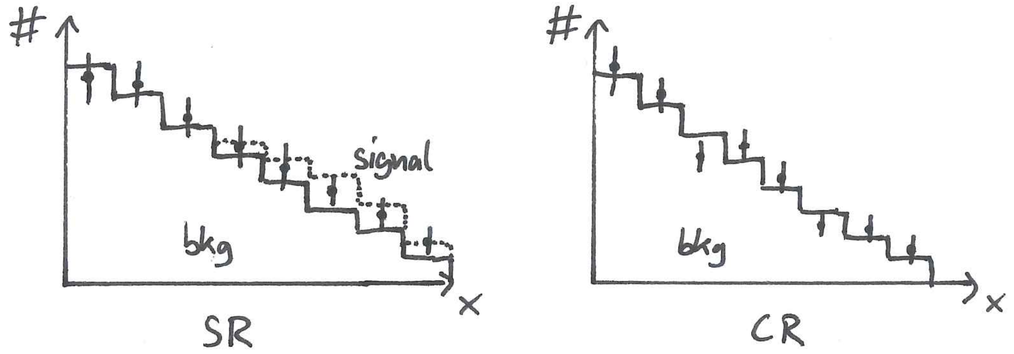
Statistical Methods & Confidence Intervals

Goal of these lectures:

- derive confidence interval for signal strength μ in this example:



- two regions (SR, CR) each with observed # events n_i and expected #events ν_i (from MC) per bin

Lecture 1

Probability:

- defined by Kolmogorov axioms:

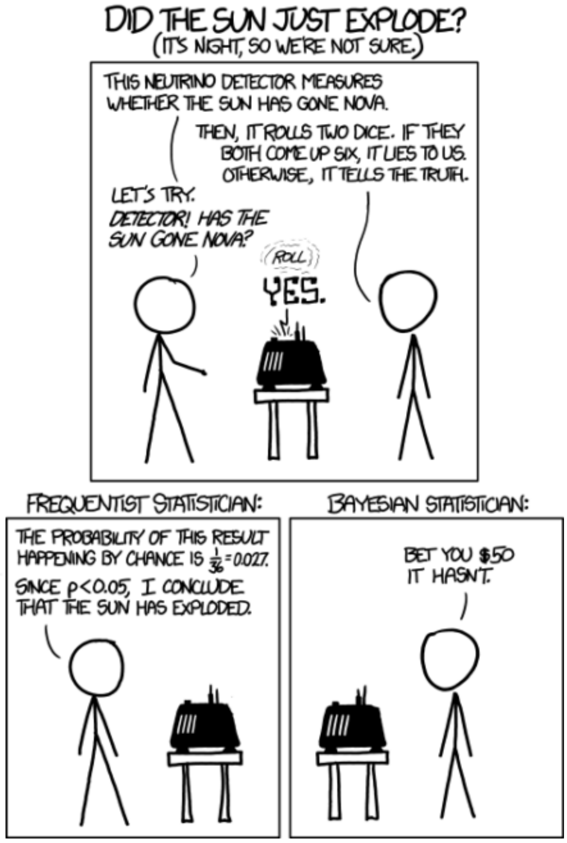
1) $P(A) \geq 0$ for any event A
2) $P(\Omega) = 1$ Ω: samplespace (all possible outcomes)
3) $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

- different interpretations of such a probability:

	Frequentist (aka. Classical)	Bayesian
	A: result of repeatable experiment $P(A) = \lim_{n\to\infty} \frac{\#A}{n}$	A: hypothesis $P(A) =$ degree of belief that A is true
caveats:	• $n\to\infty$ impossible • non-repeatable experiments? • $P(A)$ not intrinsic to A but property of ensemble	• highly subjective • how to quantify?

"Bayesians address the questions everyone is interested in by using assumptions that no one believes. Frequentists use impeccable logic to deal with an issue that is of no interest to anyone."
Louis Lyons

Bayes theorem:



- follows from definition of conditional probability

$$P(A|B)\cdot P(B) \equiv P(A\cap B) \equiv P(B|A)\cdot P(A)$$

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

- for our example:

  $A \to$ hypothesis H (with parameter μ)
  $B \to$ observation x
  $P \to$ probability density p (μ continuous)

$$\underbrace{p(H|x)}_{\text{posterior}} = \frac{\overbrace{p(x|H)}^{\text{likelihood}}\cdot \overbrace{p(H)}^{\text{prior}}}{\underbrace{\int p(x|H)\cdot p(H)\,dH}_{\text{normalization}}}$$

Bayesian: • $p(H|x)\,dH$ is the probability that H is true

Frequentist: • H is either true or not, no probability assigned
• $p(H|x)\,dH$ is the ratio that H is true in infinite repititions

Likelihood:

- $\mathcal{L}(H|x) \equiv p(x|H)$ (sometimes also written $\mathcal{L}(H)$ or $\mathcal{L}(x|H)$)

  where p: fct. of observation x for fixed H

  $$\int p(x|H)\,dx = 1$$

  $\mathcal{L}$: fct. of H for fixed observation x

  $$\int \mathcal{L}(H|x)\,dH \neq 1$$

- exact shape will depend on specific problem, but typically written as joint PDF of several standard PDFs

	PDF	EV ($\int x p(x) dx$)	variance ($\int (x-\mu)^2 p(x) dx$)
Gauss	$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	μ	σ^2
multivariate Gauss	$p(\vec{x}) \propto \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T V^{-1}(\vec{x}-\vec{\mu})\right]$	$\vec{\mu}$	V
Poisson	$p(k) = \frac{\mu^k}{k!} e^{-\mu}$	μ	μ
Binomial	$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Uniform	$p(x) = \begin{cases} \frac{1}{b-a} & , a \le x \le b \\ 0 & , \text{else} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponential	$p(x) = \begin{cases} \frac{1}{\tau}\exp(-x\tau) & , x \ge 0 \\ 0 & , \text{else} \end{cases}$	τ	τ^2
Cauchy / Breit-Wigner	$p(x) \propto \frac{\Gamma}{(x-\mu)^2 + (\Gamma/2)^2}$	undef	undef

- non-converging integrals $\int x p(x) dx$ and $\int x^2 p(x) dx$
- Can be approximated Gaussian close to peak if $\Gamma \ll \mu$

for our example:

- individual bins are independent counting experiments $\Rightarrow$ joint PDF factorizes:

$$p(\vec{n}|\vec{\nu}) = \prod_i p(n_i|\nu_i)$$

- What about systematics?

$\rightarrow$ parametrize them via nuisance parameters θ

$$\vec{\nu} = \vec{\nu}(\mu, \vec{\theta})$$

$$\mathcal{L}_{SR} = \mathcal{L}_{SR}(\mu, \vec{\theta}|\vec{n}) = \prod_{i \in SR} p(n_i|\nu_i(\mu,\vec{\theta})) \prod_{j \in syst} \underbrace{q_j(\theta_j|\overset{\text{nominal value}}{\tilde{\theta}_j})}_{\text{prior of systematic}}$$

$$\mathcal{L}_{CR} = -\text{''}-$$

$$\mathcal{L}_{tot}(\mu, \vec{\theta}|\vec{n}) = \prod_{i \in SR, CR} p(n_i|\nu_i(\mu,\vec{\theta})) \prod_{j \in syst} q_j(\theta_j|\tilde{\theta}_j)$$

$\rightarrow$ this is identical to constructing the posterior $p(\mu, \vec{\theta}|\vec{n})$ via Bayes theorem with flat prior for μ (neglecting the normalization)

$\rightarrow$ we can combine different measurements by multiplying their likelihoods (careful not to count syst. priors twice!)

- What is $p(n_i | \nu_i)$?
  - typically each bin is counting exp. → Poissonian
  - if any post-processing applied to observed spectrum (e.g. unfolding), p in general no longer Poissonian
    - → can often be approximated Gaussian in this case
    - → if bin-to-bin correlations introduced, use multivariate Gaussian or diagonalize
    - → can also use histogrammed PDF obtained via bootstrapping
- What is $q_j(\theta_j | \tilde{\theta}_j)$?
  - → depends on systematic!
    - detector syst.
      - typically considered Gaussian
      - typically parametrized s.t. $\tilde{\theta} = 0$ and variance $= 1$

$$\Rightarrow \nu_i(\vec{\theta}) = \nu_i(\tilde{\vec{\theta}}) + \sum_j \theta_j \cdot \underbrace{S_j(\nu_i)}_{\substack{\text{impact of } \theta_j \text{ on } \nu_i \\ \text{estimated by varying it } \pm 1\sigma}}$$

e.g. $q_{JES} = \mathcal{N}(0,1)$

- $q_{renorm} = \begin{cases} 1/2 & -1 \leq \theta \leq 1 \\ 0 & \text{else} \end{cases}$
- $q_{scalefac} \propto 1$
- $q_{MCstat,i} = \mathcal{N}(0,1)$
- $q_{BR(H \to bb)} = \mathcal{N}(0.58, 0.02)$

⋮