



統計誤差と上限値

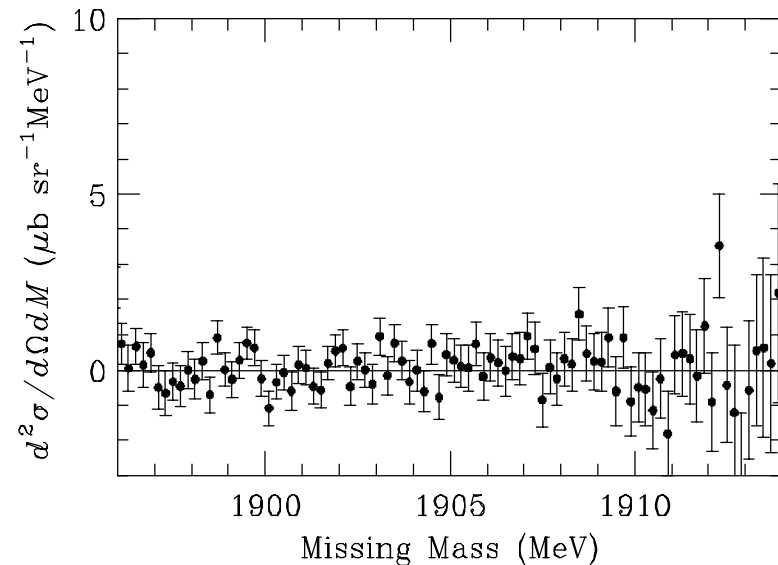
A. Tamii

今回の話の主題

- ダイバリオン探索実験の結果、生成断面積のデータが右図のような統計誤差付きで得られた。

この結果から、生成断面積の上限値を求めよう!

ダイバリオン探索実験結果



- 例題 (物理境界付近の統計誤差)
 - 100個の不安定核を生成した。ある崩壊チャンネルを検出効率10% で測定すると9 イベントを得た。そのチャンネルへの分岐比は何% か。

$$90 \pm 30\% ?$$

- 偏極分解能 0.5 の偏極度計で陽子ビームの偏極度を測定した。左散乱 400 個、右散乱 120 個が測定された。ビーム偏極度を求めよ。

$$1.077 \pm 0.074 ?$$



目次

- 巷 (ちまた) の統計
- 記法、代表的な確率分布
- 真の値の推定とベイズ原理
- 物理的境界付近の統計誤差 (例 1)
- 上限値の決定
- まとめとコメント
- 参考文献



巷 (ちまた) の統計

- 横断歩道での交通事故記録の統計をとったところ、青信号の横断中の方が、赤信号の横断中よりも、交通死亡事故が多いという統計結果が得られた。

~~青信号で渡った方が事故にあう危険が高い。~~

青信号で渡る人の方が圧倒的に多いため

母集団における分布をとりしれていない場合

- 肺癌による死亡者の統計をとったところ、郊外で空気のよい地域ほど人口に対する死亡率が高いという結果が得られた。

~~空気のよい地域のほうが肺癌になる可能性が高い。~~

郊外で静養中の患者が死亡すると死亡率が高くなる

相関はあるが、因果関係が逆である場合



巷 (ちまた) の統計

- 統計誤差が考慮されていない場合
例：被験者 1 人の実験で結論を出してしまうテレビ番組の検証実験
- 標本が無作為抽出されていない場合
例：インターネットによる視聴者からのアンケート調査の統計
- 標本にバイアスをかけてしまう場合。
例：不適切な心理学実験など
- 作為的なデータの棄却、選択。
例：成功した例しか掲載しないダイエット広告

我々は本当に正しく統計を扱っていると言えるのだろうか？



使用する記法

- 確率変数 (random variable)
 X, Y, Z, \dots
- 標本値 (sample value) または観測値 (observation value)
 $x, y, z, \dots, \hat{\mathbf{S}}, \dots$
- 真の値、母数 (parameter)
 \mathbf{m}, \dots
- 確率密度関数 (Probability Density Function: p.d.f.)
 $f(x), f(\mathbf{m})$
- 条件付き確率密度関数 (Conditional p.d.f.)
 $f(x|\mathbf{m}), f(\mathbf{m}|x)$
- $f(x|\mathbf{m})$ が確率分布 $B_N(n, \mathbf{m})$ に従う
 $f(x|\mathbf{m}) \sim B_N(n, \mathbf{m})$
- x X の平均値 $E\{X\}$, \bar{X} 、分散 $V\{X\}$

代表的な確率分布(離散分布)

2項(Binomial)分布 $B_N(n, p)$

“独立試行”

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$E\{X\} = np, \quad V\{X\} = np(1-p)$$

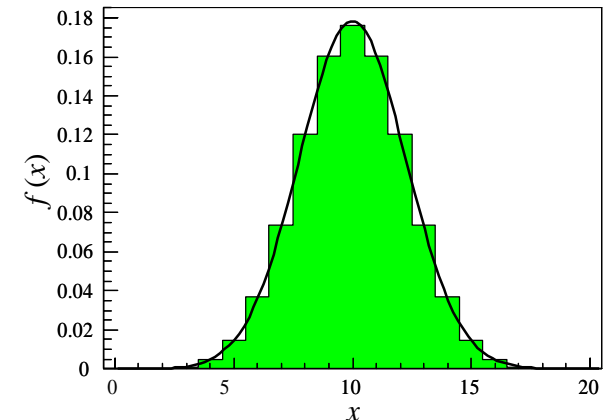
- n 大のとき正規分布に近づく

$$\lim_{n \rightarrow \infty} B_N(n, p) \sim N(np, np(1-p))$$

- $np = \text{一定}$ で、 $p \rightarrow 0$ の時、Poisson 分布に近づく。“希現象の大量観測”

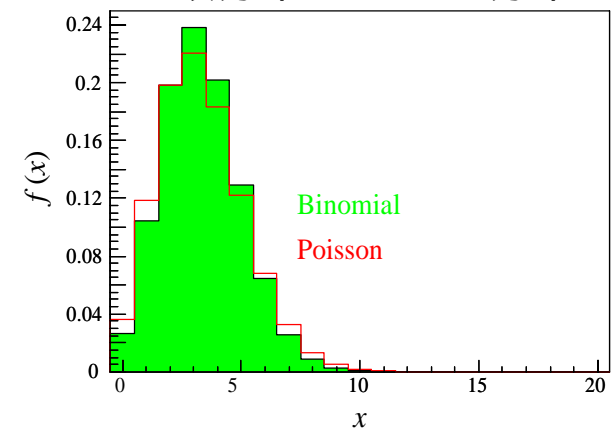
$$\lim_{p \rightarrow 0} B_N(n, p) \Big|_{np=\text{const}} \sim P_0(np)$$

2項分布と正規分布



例: コインを20回投げた。表は何回出るか。実践は正規分布近似。

2項分布とPoisson 分布



例: サイコロを20回投げた。1は何回出るか。Poisson 分布近似とは、少しずれがみられる。

代表的な確率分布(離散分布)

Poisson 分布 $P_0(I)$

“独立試行”、“希現象の大量観測”

$$f(x|I) = \frac{I^x}{x!} \exp(-I)$$

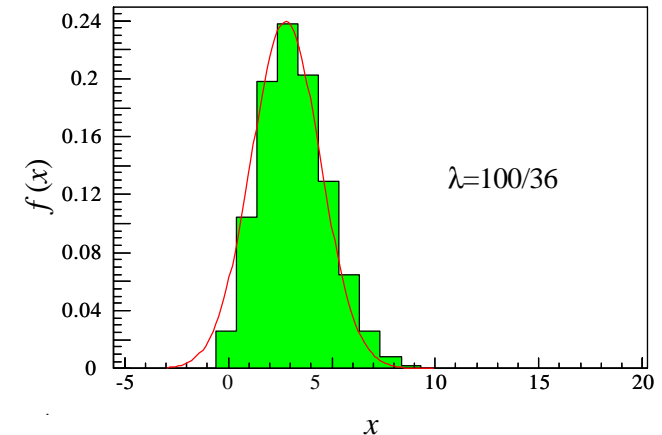
$$E\{X\} = I, V\{X\} = I$$

(計数測定の統計誤差 \sqrt{N} の起源)

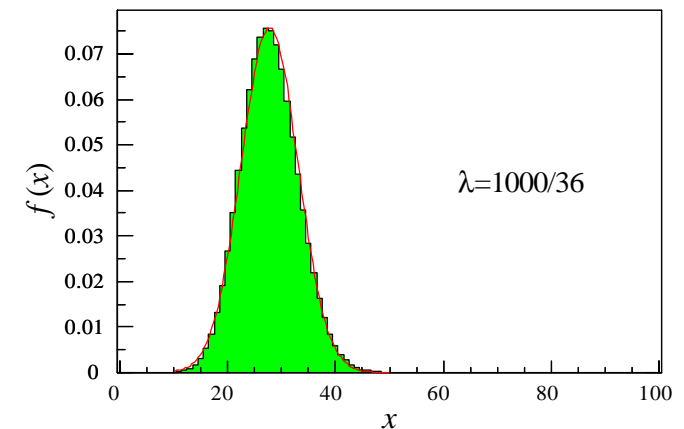
- I 大のとき正規分布に近づく

$$\lim_{I \rightarrow \infty} P_0(I) \sim N(I, I)$$

Poisson 分布と正規分布近似



例：2個のサイコロを100回投げると、何回ピンゾロ(1・1)が出るか。



例：2個のサイコロを1000回投げると、何回ピンゾロが出るか。 8



代表的な確率分布(離散分布)

超幾何(Hypergeometric)分布

例：トランプからカードを4枚同時に引いた。
4枚の中にスペードが2枚ある確率は？

幾何(Geometric)分布

例：サイコロで1の目がでるまでに、何回振りつづける必要があるか？

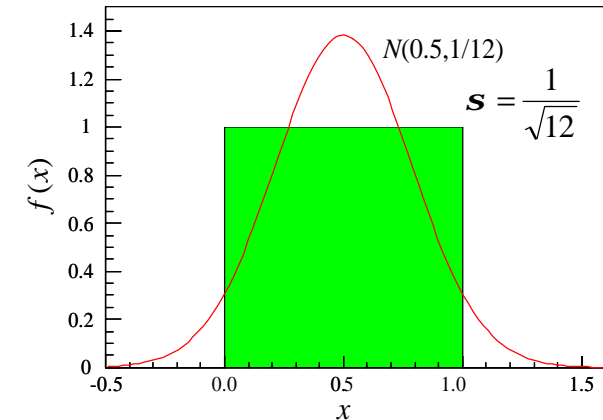
代表的な確率分布(連続分布)

一様分布と正規分布近似

一様(Uniform)分布 $U(\mathbf{a}, \mathbf{b})$

$$f(x | \mathbf{a}, \mathbf{b}) = \frac{1}{\mathbf{b} - \mathbf{a}}$$

$$E\{X\} = \frac{\mathbf{a} + \mathbf{b}}{2}, \quad V\{X\} = \frac{(\mathbf{a} - \mathbf{b})^2}{12}$$



正規(Normal)分布、ガウス(Gaussian)分布

$$f(x | \mathbf{m}, \mathbf{s}^2) = \frac{1}{\sqrt{2\pi\mathbf{s}^2}} \exp\left(-\frac{(x - \mathbf{m})^2}{2\mathbf{s}^2}\right)$$

$N(\mathbf{m}, \mathbf{s}^2)$

$$E\{X\} = \mathbf{m}, \quad V\{X\} = \mathbf{s}^2$$

$$\text{FWHM} = 2\sqrt{2\ln 2} \times \mathbf{s} = 2.35\mathbf{s}$$

- 正規分布 $N(\mathbf{m}, \sigma^2)$ の n 個の標本平均 \bar{X}_n は、分散 \mathbf{s}^2/n の正規分布に従う

$$f(x) \sim N(\mathbf{m}, \mathbf{s}^2) \quad \Rightarrow \quad f(\bar{X}_n) \sim N(\mathbf{m}, \mathbf{s}^2/n)$$

代表的な確率分布(連続分布)

Student's t 分布

正規分布である母集団の分散 (σ^2) が未知の場合の標本分布。
母平均 m 標本数 n 、標本分散 \hat{s}^2 の時、 X の T -変換は t_n 分布に従う

$$T = \frac{X - m}{\hat{S}} \sim t_{n-1}$$
$$t_n = f_n(t) = \frac{1}{\sqrt{np}} \frac{\Gamma[(n+1)/2]}{\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

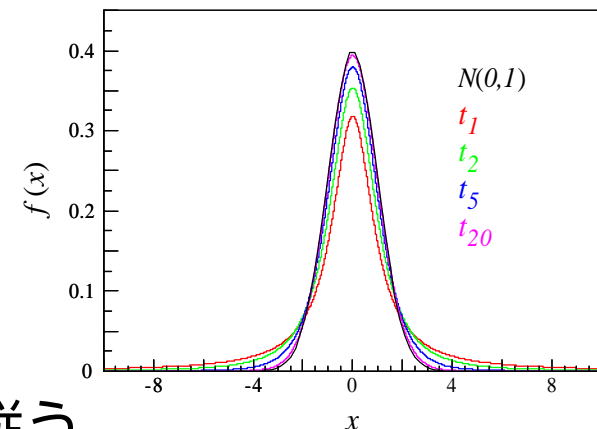
- 標本平均 \bar{X}_n の不偏標本分散を $\hat{s}^2 = \frac{n}{n-1} \hat{S}^2$ とおくと 以下の \bar{X}_n の T -変換は t_{n-1} 分布に従う

$$T = \frac{\bar{X}_n - m}{\hat{s}/\sqrt{n}} \sim t_{n-1}$$

- n 大のとき t_n は正規分布に近づく。

$$\lim_{n \rightarrow \infty} t_n \sim N(0,1)$$

Student's t 分布と正規分布





代表的な確率分布(連続分布)

指数(Exponential)分布

例：素粒子が崩壊までにかかる時間

Gamma 分布

例：正規分布の χ^2 の分布関数

Beta 分布



真の値の推定 (例題 1)

問) 新種の不安定核の測定に成功した。解析の結果、得られた不安定核イベントは5個で、バックグラウンドイベントは無視できる程度であることが分かった。

不安定核の生成断面積が 1 nb/sr の場合1イベント測定されるような実験条件 (ビーム量、標的厚、立体角、検出効率等々) であった時、この不安定核の生成断面積を統計誤差付きで述べよ。

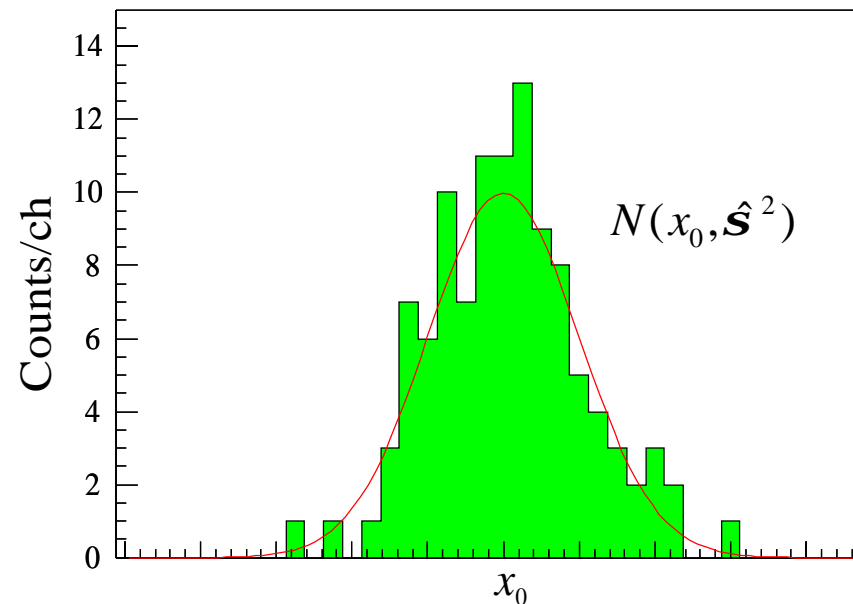
答) 不安定核の生成個数は Poisson 分布に従うので、生成個数は $5 \pm \sqrt{5} = 5.0 \pm 2.2$ 個。

よって、不安定核の生成断面積は $5.0 \pm 2.2 \text{ nb/sr}$ である。

(測定個数が 0 個の場合はどうするの?)

真の値の推定 (例題 2)

問) m の値 (例えば励起状態のエネルギー) を測定する実験を行い、100 個のデータを取得した。 m の測定値 x の分布として、下図のような正規分布 $N(x_0, \hat{s}^2)$ に近い分布が得られた。
真の値 m はどのように推定されるか。



m の測定値 x の分布



大数の法則と不偏有効推定量

大数の法則 (law of large numbers) :

母平均 m 母分散 s^2 の同一分布から得られた、独立な n 個の標本の平均を \bar{X}_n と置くと

$$E\{\bar{X}_n\} = m, \quad V\{\bar{X}_n\} = \frac{s^2}{n}$$

よって、標本数 n を大きくするとき、標本平均 \bar{X} は母平均 m に収束する。

m の測定値 x が正規分布をとる場合、 x の平均値 \bar{X} は、 m の不偏(unbiased) 有効(efficient) 推定量(estimator)である。

不偏 : 同様の測定を何度も行った場合、推定量は真の値の周りに均等に分布する。

有効 : 与えられた情報量を最大限に生かした推定であることを意味する。

真の値の推定 (例題 2)

答) 母分散が未知であるため、確率密度分布として Student's t 分布となるが、計数 $n (=100)$ が大きいので正規分布で近似する。

真の値 m の推定量として標本平均 \bar{x} を用いると、 m の確率密度分布として $N(x_0, \hat{s}^2 / (n-1))$ が得られる。

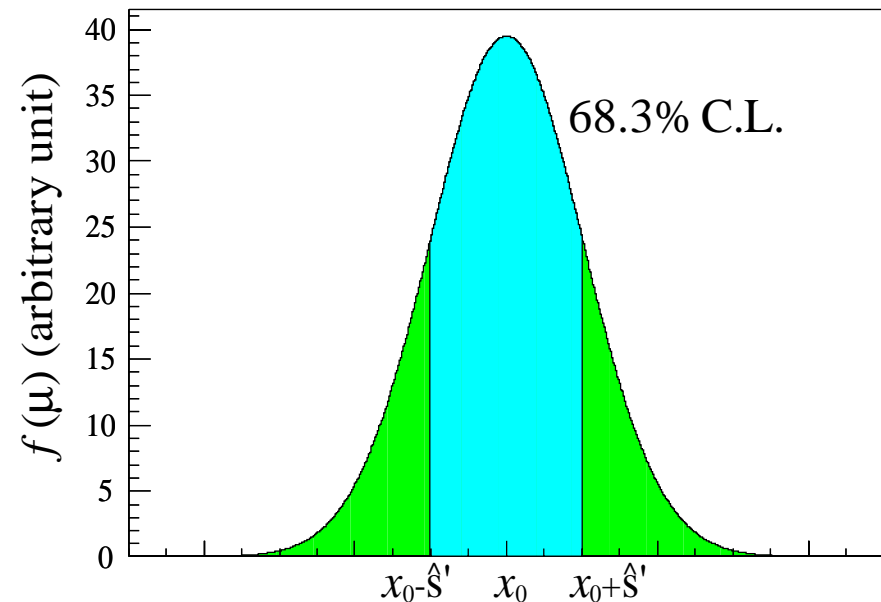
$$(\hat{s}')^2 = s^2 / (n-1)$$

とおくと

$$m = x_0 \pm \hat{s}' \quad \text{at 68.3\% C.L.}$$

$$m = x_0 \pm 2\hat{s}' \quad \text{at 95.4\% C.L.}$$

$$m = x_0 \pm 3\hat{s}' \quad \text{at 99.7\% C.L.}$$



m の確率密度分布



真の値の推定 (狂牛病の牛)

問) 狂牛病の検査試験の感度が

狂牛病に罹っている場合 ... 100.0% 陽性

狂牛病に罹っていない場合 ... 99.8% 陰性、0.2%陽性

であるとする。

あなたの飼っている牛 1頭の検査を行ったところ、陽性反応が出た。
その牛が狂牛病に罹っている可能性は何%か。

答) 陽性反応が出たとしても、狂牛病に罹っていない可能性がある。

狂牛病に罹っている場合、陽性が出る可能性は 100%、

狂牛病に罹っていない場合、陽性が出る可能性は 0.2%。

よって、狂牛病に罹っている可能性は 99.8% であろう

$$\frac{1.000}{1.000 + 0.002} \sim 0.998$$



事前確率の仮定

- ところが、注意深く考えてみると、前ページの議論では、検査前の時点における狂牛病に罹っている確率と罹っていない確率が1対1であることが暗黙のうちに仮定されている、ということが分かる。

例えば、全国調査で、狂牛病に罹っている牛の割合が、10万頭に対して1頭 (10^{-5}) であることが分かっていた場合、陽性反応が出た牛が狂牛病に罹っている確率は0.5% である。

$$\frac{10^{-5} \times 1.000}{10^{-5} \times 1.000 + (1 - 10^{-5}) \times 0.002} \cong 0.005$$

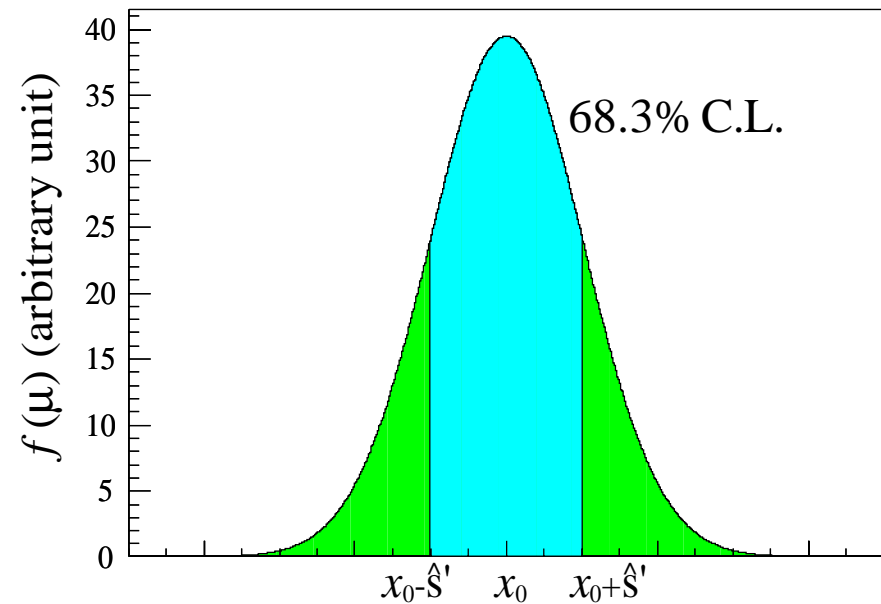
検査以前の確率分布 (事前確率) が、検査後の確率分布 (事後確率) に影響する。

事前確率を仮定しなければ事後確率を決定することができない。

真の値の推定 (例題 2) 再考

図のような真の値 m の確率分布は、 m の事前分布が対象となる領域において一様分布であった場合にのみ、正しい答えである。

例えば、物理的に $x_0 - \hat{s}'$ 以下の m の値はありえない (素粒子の質量測定において質量が負になるなど) 場合には、図の確率分布は正しい答えではない。



m の確率密度分布



ベイズ原理 (Bayes' Principle)

- x, \mathbf{m} の同時密度関数 (joint density function) は

$$f(x, \mathbf{m}) = f(\mathbf{m} | x) f(x) = f(x | \mathbf{m}) f(\mathbf{m})$$

と表される。これを变形して、

$$\begin{aligned} f(\mathbf{m} | x) &= \frac{f(x | \mathbf{m}) f(\mathbf{m})}{f(x)} \\ &= \frac{f(x | \mathbf{m}) f(\mathbf{m})}{\int_{\mathbf{m}'} f(x | \mathbf{m}') f(\mathbf{m}') d\mathbf{m}'} \propto f(x | \mathbf{m}) f(\mathbf{m}) \end{aligned}$$

この式により、事前確率 (prior probability) $f(\mathbf{m})$ から事後確率 (posterior probability) $f(\mathbf{m}|x)$ が得られる。

分母の値は規格化の為に必要。

離散分布の場合、分母の積分は和 (Σ) に置き換わる。

真の値の推定 (例題 1) 再考

問) 新種の不安定核 $x=5$ 個の測定に成功した...(中略)... 生成断面積を統計誤差付きで述べよ。

答) 不安定核の生成個数は Poisson 分布に従うので、

$$f(x | \mathbf{m}) \sim P_0(\mathbf{m}) = \frac{\mathbf{m}^x}{x!} \exp(-\mathbf{m})$$

事前確率として 0 から無限大までの一様分布 $f(\mathbf{m})=U(0, \infty)$ を仮定し、Bayes 原理を適用すると、生成個数の事後確率は、

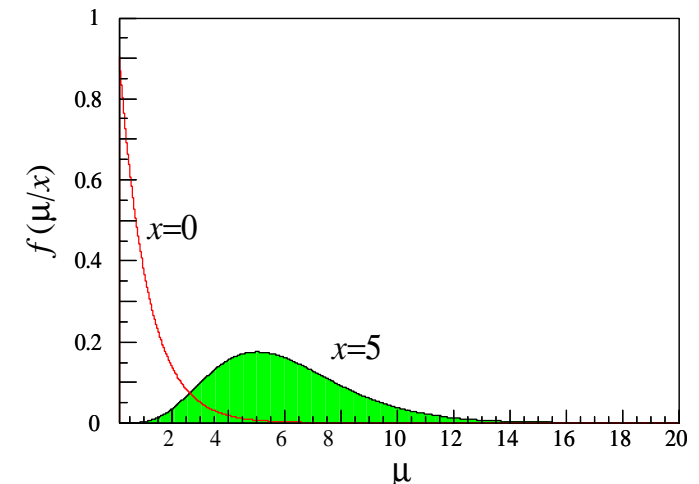
$$\begin{aligned} f(\mathbf{m} | x) &= \frac{f(x | \mathbf{m}) f(\mathbf{m})}{\int_0^\infty f(x | \mathbf{m}') f(\mathbf{m}') d\mathbf{m}'} = \frac{f(x | \mathbf{m})}{\int_0^\infty f(x | \mathbf{m}') d\mathbf{m}'} \\ &= f(x | \mathbf{m}) = \frac{\mathbf{m}^x}{x!} \exp(-\mathbf{m}) \end{aligned}$$

$$E_m\{f(\mathbf{m} | x)\} = V_m\{f(\mathbf{m} | x)\} = x + 1$$

は $6 \pm \sqrt{6} = 6.0 \pm 2.4$ 個。

よって、不安定核の生成断面積は 6.0 ± 2.4 nb/sr。ただし、生成断面積の最頻値は 5.0 nb/sr である。

生成個数の確率密度分布





物理境界付近の統計誤差 (例題1) 1/4

100個の不安定核を生成した。ある崩壊チャンネルを検出効率10% で測定すると9 イベントを得た。そのチャンネルへの分岐比は何% か。

- [Poisson 分布による近似計算]

最も単純な計算では、統計誤差は Poisson 分布に従うとして、
測定個数 ... 9 ± 3 個

検出効率の10%で割って、
測定チャンネルへの崩壊個数 ... 90 ± 30 個

崩壊前の数が100個であるので、
分岐比 ... $90 \pm 30\%$



物理境界付近の統計誤差 (例題1) 2/4

- [2項分布による計算]

崩壊過程と検出過程は正確には2項分布に従う

崩壊前の数を n 、測定チャンネルへの分岐比を m 、崩壊個数を y 、検出効率を e と置くと

$$f(y | \mathbf{m}, n) \sim B_N(n, \mathbf{m}), \quad f(x | \mathbf{e}, y) \sim B_N(y, \mathbf{e})$$

よって測定個数 x の分布関数は

$$f(x | \mathbf{m}, \mathbf{e}, n) = \sum_{y=x}^n f(x | \mathbf{e}, y) f(y | \mathbf{m}, n) = \binom{n}{x} (\mathbf{m}\mathbf{e})^x (1 - \mathbf{m}\mathbf{e})^{n-x} \sim B_N(n, \mathbf{m}\mathbf{e})$$

となり、確率 $\mathbf{m}\mathbf{e}$ の2項分布に従う(物理的に納得できる)。

測定個数の統計誤差を2項分布の分散の式 $V\{X\} = np(1-p)$ から求めると、 $p = \mathbf{m}\mathbf{e} = 0.09$ であるので分散は 8.19、標準偏差は 2.86。

測定個数は 9 ± 2.86 個

分岐比 ... **90.0 ± 28.6%**

前ページの計算は、 $\mathbf{m}\mathbf{e} = 0.09$ が 1 より十分小さいとして、2項分布を Poisson 分布に近似した場合に相当する。大きな差はない。

物理境界付近の統計誤差 (例題1) 3/4

- [Bayes 原理 による推定]

事前確率として $U(0,1)$ の一様分布を仮定し
 $f(x|\mathbf{m})$ として2項分布を取る。

$$f(\mathbf{m}) \sim U(0,1)$$

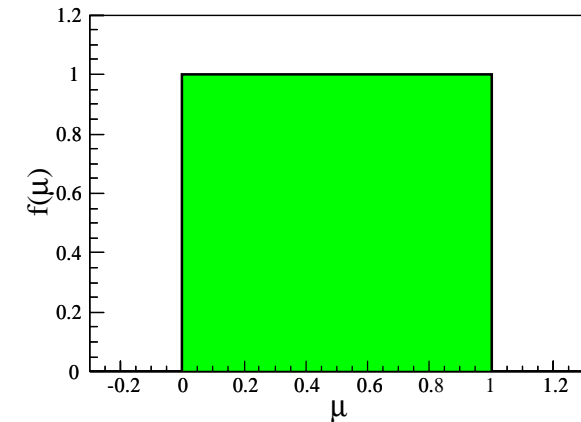
$$f(x|\mathbf{m}, \mathbf{e}, n) = \binom{N}{x} (\mathbf{m}\mathbf{e})^x (1-\mathbf{m}\mathbf{e})^{N-x}$$

Bayes 原理を適用すると 事後確率は

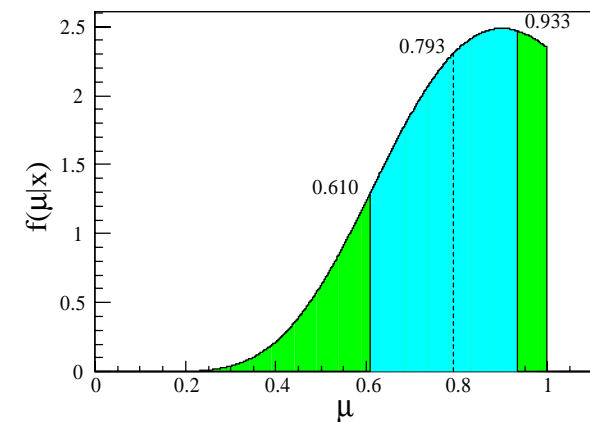
$$f(\mathbf{m}|x, \mathbf{e}, n) = \frac{f(x|\mathbf{m}, \mathbf{e}, n)f(\mathbf{m})}{\int_0^1 f(x|\mathbf{m}', \mathbf{e}, n)f(\mathbf{m}')d\mathbf{m}'}$$

右図のような確率密度が得られる。

m の事前確率分布



m の事後確率分布



物理境界付近の統計誤差 (例題1) 4/4

- 得られた確率密度分布をそのまま論文に掲載する訳にも行かないので、分布から統計量を抽出することになるが、抽出方法には任意性がある。

- 平均値と標準偏差から算出した分岐比

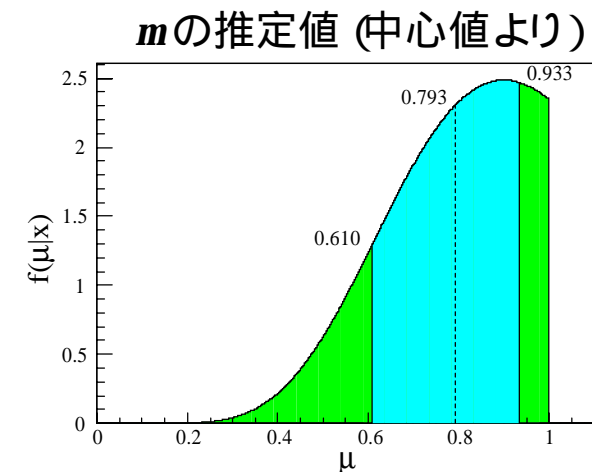
$$77.3 \pm 15.1\%$$

- 中心値と68.26%の信頼区間から算出した分岐比

$$79.3^{+14.0}_{-18.3} \%$$

2 σ ずれると100%を超えることになるが、これは図のような確率密度分布から統計量を無理やり引き出したことに原因がある。

確率密度分布自身は、正しく0から1の間に収まっている。



最尤法 (Maximum Likelihood Method) による推定

- 最尤法では、測定値の同時密度関数 (≒尤度関数)

$$f(x | \mathbf{m}) = \prod_{i=1}^n f(x_i | \mathbf{m})$$

が最大になる値を**最尤推定値**(Maximum Likelihood Estimator)とする。
今の例題の場合、得られた密度分布は尤度関数に比例しており
右図より 最尤法による推定値は 0.900 である。

統計誤差を求める方法には任意性があるが、正規分布近似による分散の式

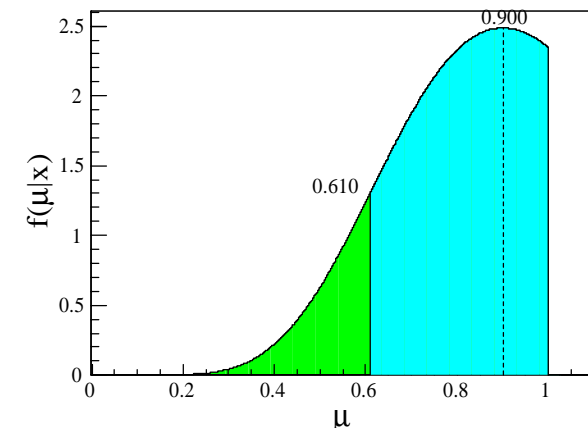
$$V\{X\} = -\left(\frac{\partial^2 \log L}{\partial^2 x}\right)_{x=E\{V\}}^{-1}$$

から求めた場合には 0.286。

よって分岐比は **90.0 ± 28.6%** となる。

なお、**最尤法では一様な事前確率を暗黙の内に仮定していることに注意。**

最尤法による推定





事前確率をどう決めるべきか

- Bayes 原理を適用するには、事前確率が必要である。
(Bayes原理以外の推定方法では、事前確率は暗黙のうちに仮定されている)
- 事前確率の決め方については統一の見解はない。
(統計学は事前確率の決定法については何も教えてくれない)
- 一般的には、事前確率として一様分布がとられる場合が多い。
しかしこの場合でも、何に関して一様かという不定性が残る。
例えば、素粒子の質量(m)を決める場合、 m 、 m^2 、 m^{-1} 等の何に関して一様分布を取ればよいのか、というのは必ずしも定かではない。

物理的直感に照らして適切と思われる分布をとる (より他ない!)

物理境界付近の統計誤差 (例題2)

- 偏極分解能 $A_y = 0.5$ の偏極度計で陽子ビームの偏極度を測定した。左散乱 400 個、右散乱 120 個が測定された。ビーム偏極度 p_y を求めよ。

- 誤差伝播の式から、

$$p_y = \frac{1}{A_y} \frac{L-R}{L+R} \quad \Delta p_y = \frac{1}{A_y} \frac{2LR}{L+R} \sqrt{\left(\frac{\Delta L}{L}\right)^2 + \left(\frac{\Delta R}{R}\right)^2}$$

左散乱(L)、右散乱(R)の計測数がそれぞれ Poisson 分布に従うとすると $\Delta L = \sqrt{L}$, $\Delta R = \sqrt{R}$ より、

$$p_y = 1.077 \pm 0.074$$

推定値が物理的に許容される範囲外に出ている。この計算では事前確率として $U(-, +)$ が暗黙のうちに仮定されているためである。

物理境界付近の統計誤差 (例題2)

- L と R の比 r は p_y の関数であり 次の式で表される。

$$r \equiv \frac{R}{L} = \frac{1 - p_y A_y}{1 + p_y A_y}$$

L, R は Poisson 分布に従うと仮定し、正規分布で近似する。 L, R の同時密度関数は、 p_y と L の期待値 L_0 によって表現され (R の期待値は rL_0 となる)

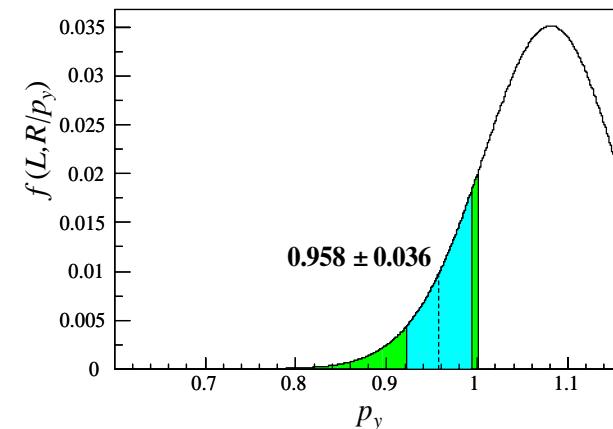
$$f(L, R | L_0, p) = f(L | L_0, p_y) f(R | rL_0, p_y) = \frac{1}{2p} \frac{1}{L_0 \sqrt{r}} \exp\left(-\frac{(L-L_0)^2}{2L_0} - \frac{(R-rL_0)^2}{2rL_0}\right)$$

事前確率として、 p_y は $U(-1,1)$ の一様分布、 L_0 は $U(0, \quad)$ の一様分布を仮定すると

$$f(L, R | p_y) \propto \frac{1}{\sqrt{r}} \int_0^\infty dL_0 \frac{1}{L_0} \exp\left(-\frac{(L-L_0)^2}{2L_0} - \frac{(R-rL_0)^2}{2rL_0}\right)$$

この式から、右図の様な確率密度分布が得られる。
(なお、 L_0 を 400 に固定して計算してもほぼ同じ結果が得られる)
この分布の平均値と分散から求められる偏極度は

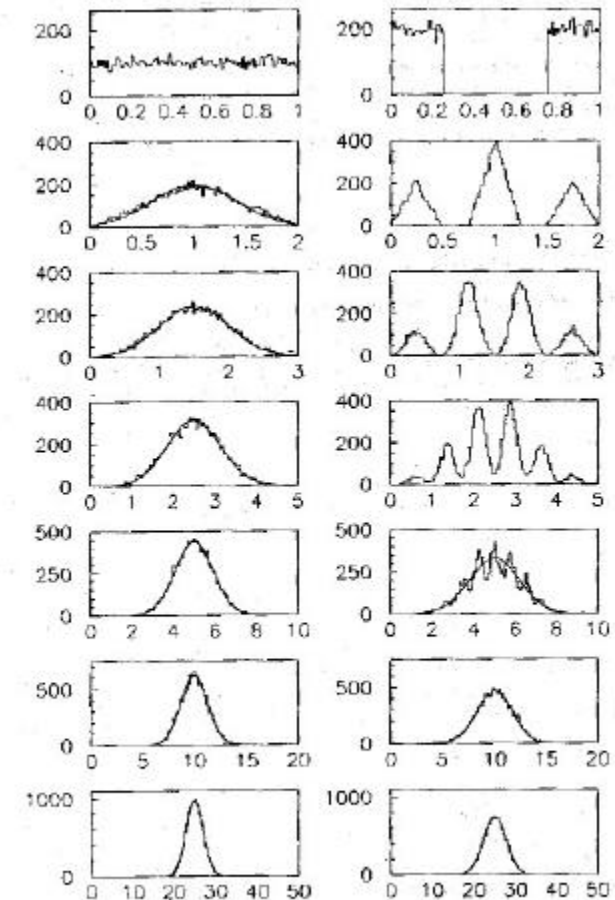
$$p_y = 0.958 \pm 0.036$$



(余談)なぜ系統誤差分布に正規分布を用いるのか？

- **中心極限定理 (Central Limit Theorem) の要点**
 X_1, X_2, \dots, X_n が独立で同一の分布に従う時、その平均値は n において、元の分布と同じ平均値、分散を持つ**正規分布に収束する**。
これはいかなる母分布に関しても成り立つ。

正確な定理と証明については、数理統計学の教科書 (例えば Ref. [1]) を参考のこと。

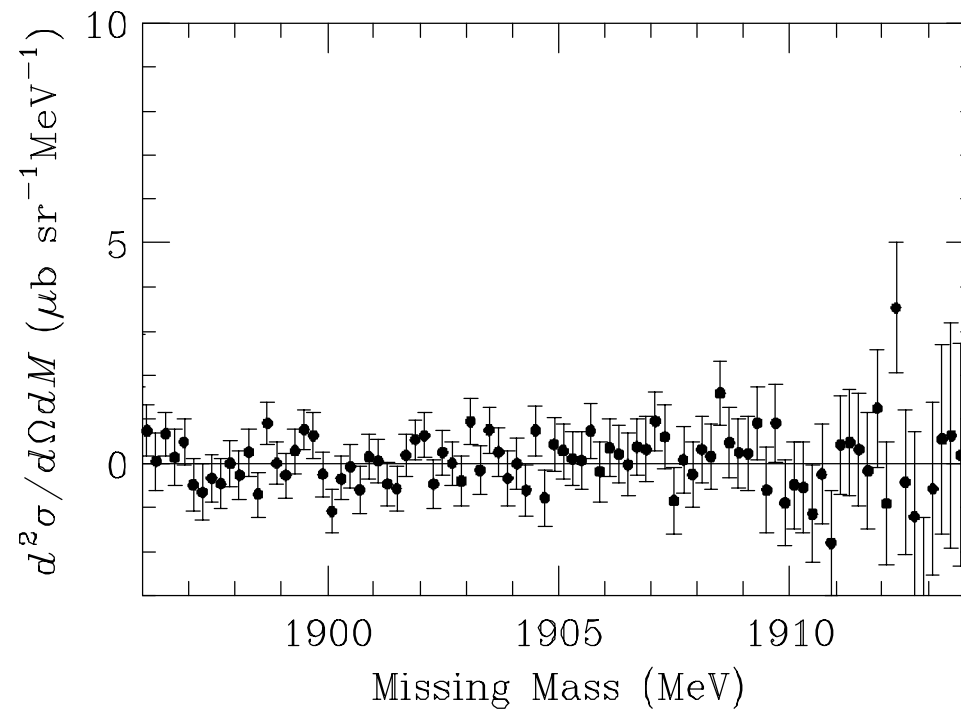


3: Central limit theorem at work: The sum of n variables, for two different distributions, as values of n (top to bottom) are 1, 2, 3, 5, 10, 20, 50.

収束の例。上から1,2,3,5,10,20,50
回の測定の平均値 (Ref. [2]) 30

上限値の計算 (例題)

- ダイバリオン探索実験を行ったところ、図のような結果が得られた。誤差は、バックグラウンドを差し引いた後の統計誤差である。ダイバリオンの幅 (FWHM) を 1 MeV と仮定し、生成断面積の上限を求めよ。



上限値の計算 (計数値自身の場合)

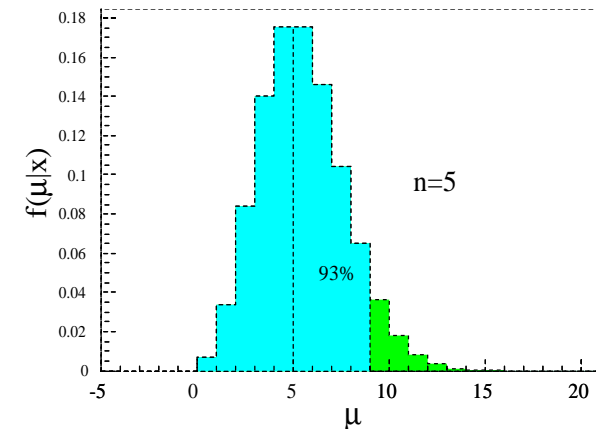
- ある1つのビンに着目する。
計数値 n の Poisson 分布を求め、信頼度に対応する数値以上 (conservative) でもっとも小さい値を取り、上限値とする。

$$f(\mathbf{m} | n) \sim P_o(n)$$

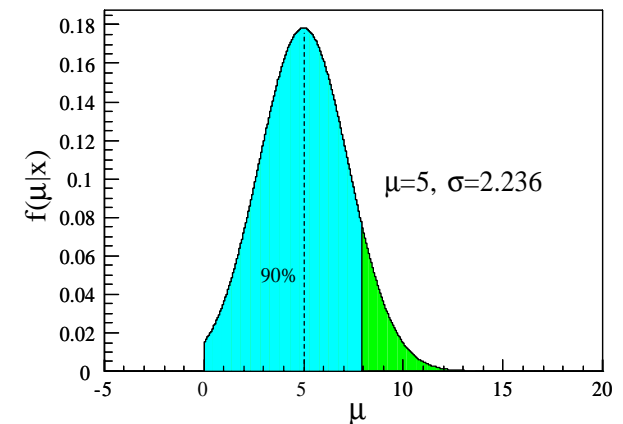
- 計数値が大きい場合には正規分布で近似できる。正規分布の0以下を排除した分布を作り、信頼度に対応する上限値を決める。

$$f(\mathbf{m} | x, \mathbf{s}) \sim \begin{cases} 0 & m < 0 \\ \mathbf{a}N(x, \mathbf{s}^2) & m \geq 0 \end{cases}$$

Poisson 分布 90% C.L.



正規分布 90% C.L.



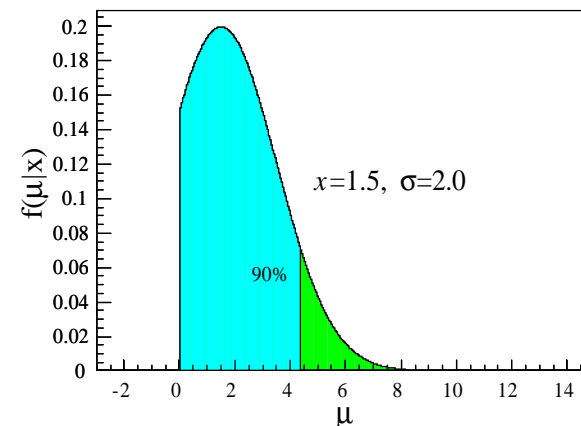
上限値の計算 (測定誤差がある場合)

- ある1つのビンに着目する。
正規分布の測定誤差がある場合、正規分布の0以下を排除した分布を作り、信頼度に対応する上限値を決める。

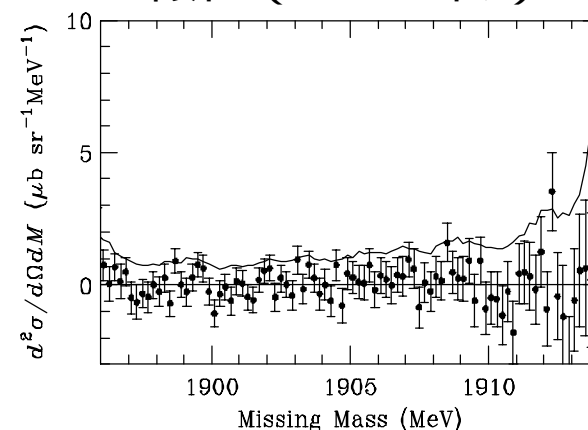
$$f(m|x,s) \sim \begin{cases} 0 & m < 0 \\ aN(x,s^2) & m \geq 0 \end{cases}$$

- 測定にバックグラウンドがある場合は、何らかの方法でバックグラウンドを見積もり (直線フィットなど)、バックグラウンドを差し引いた実験値と誤差を決める。
その後、上と同様にして上限値を決める。
各ビンにおける上限値が得られた。

正規分布 90% C.L.



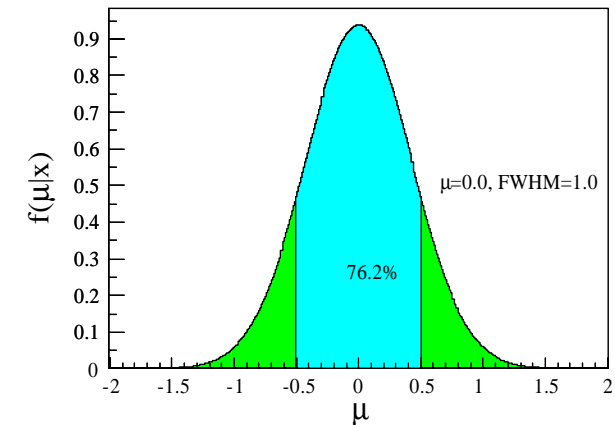
ダイバリオン生成断面積
上限値 (1 MeV 当り)



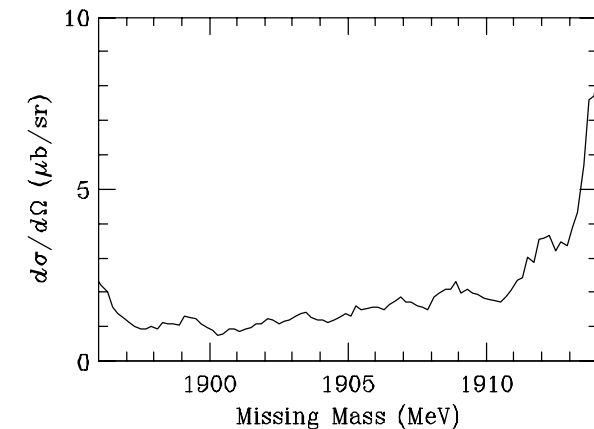
上限値の計算

- 適当なエネルギービンをとり (今の場合 1 MeV)そのビン内での上限値を決める。
- ダイバリオンのピークの幅を仮定し (今の場合FWHM で 1 MeV)、ピークの面積のうち、**ビン幅に入る部分の比率**を求める。
- **ビン当りの上限値をこの比率で割り**、ダイバリオン生成断面積の上限を得る。

エネルギービンに入る割合



生成断面積上限値





信頼区間と上限値に関するコメント

- 信頼区間と上限値の提出に関して、多くの問題が未解決である。
例えば：
 1. 信頼区間を推測値の上下にどう振り分けるべきか。
 2. 実験データが得られたとき、信頼区間を提出するべきか上限値を提出するべきかの判断基準が曖昧である。この“物理学者の態度の変化”によって、実験データにギャップが生じる。
 3. Bayes 原理を適用する場合、事前確率をどうとるのが正しいか。
- 1,2 に関する対処法の 1つが提唱されている (unified method) 。
G.J. Feldman and R.D. Cousins, Phys. Rev. D **57** (1998) 3873.
Particle Data Group, Euro. Phys. J. C **15** (2000) 1., Chap. 8.
- 信頼区間に関するワークショップなども行われている。
Workshop on Confidence Limits, CERN, 17-19 Jan. 2000,
see <http://www.cern.ch/CERN/Division/EP/Events/CLW/>.



まとめ

- 1) 真の値の推定 (事後確率の計算)には事前確率が必ず必要である。Bayes 原理以外の推定方法では、事前確率が暗黙の内に仮定されている。最尤法でも、一様な事前確率が暗黙の内に仮定されている。
- 2) 事前確率を正しく(陽に)扱って事後確率を求めるには、Bayes 原理を用いるのが良いであろう。
- 3) 事前確率として一般には一様分布が用いられるが、どのような分布を用いるべきかについては定まった見解はない。
- 4) 信頼区間の決め方に関しても、一般には、推測値の上下にどう振り分けるかは明らかではない。
- 5) 信頼区間の提出方法については、現在も多くの議論が行われている。



コメント

- データ解析の際に、最初から Bayes 原理を適用して、測定結果が物理的に可能な領域に入るように解析するのは良くない。

解析結果が物理的に無意味な領域に入るようであれば、まずは解析の不備、統計誤差の算出ミス、系統誤差の影響などを疑うべきである。

また、そのようなデータが得られる確率が、統計的にあまりにも稀ではないか、ということについても検討すべきである。

この発表資料では Bayes 原理を用いた場合の上限値や信頼区間の計算結果を紹介しているが、必ずしも Bayes 原理を用いた結果の Publish を勧めている訳ではないことにご注意ください。特に計数 n の測定結果を $(n+1) \pm \sqrt{n+1}$ として提出すべきかどうかについては多くの異論があることでしょう。平均値を提出すべきか、最頻値を提出すべきか等についても議論は決して単純ではありません。



参考文献

1. 稲垣宣生 : 数理統計学、数学シリーズ、裳華房、1990 .
2. G. D'Agostini, *Bayesian Reasoning in High-Energy Physics: Principles and Applications*, CERN 99-03, 1999.
3. B.P. Roe, *Probability and Statistics in Experimental Physics 2nd Edition*, (Springer-Verlag, New York, 2001).
4. O. Helere, Nucl. Instrum. And Meth. **212** (1983) 319.
5. Particle Data Group, Chap 28, *Statistics*, Phys. Rev. D **54** (1996) 1.
6. Particle Data Group, Chap 28, *Statistics*, Euro. Phys. J. C **15** (2000) 1.
7. G.J. Feldman and R.D. Cousins, Phys. Rev. D 57 (1998) 3873.
8. Workshop on Confidence Limits, CERN, 17-19 Jan. 2000,
see <http://www.cern.ch/CERN/Division/EP/Events/CLW/>.