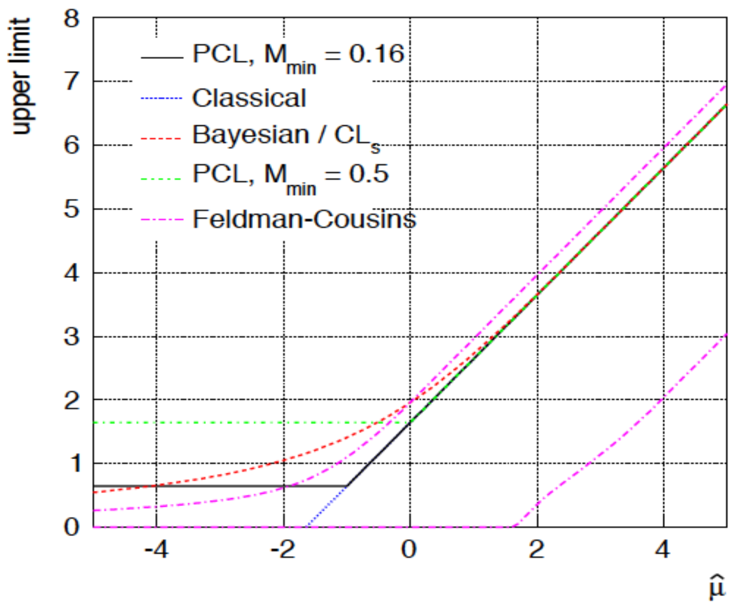
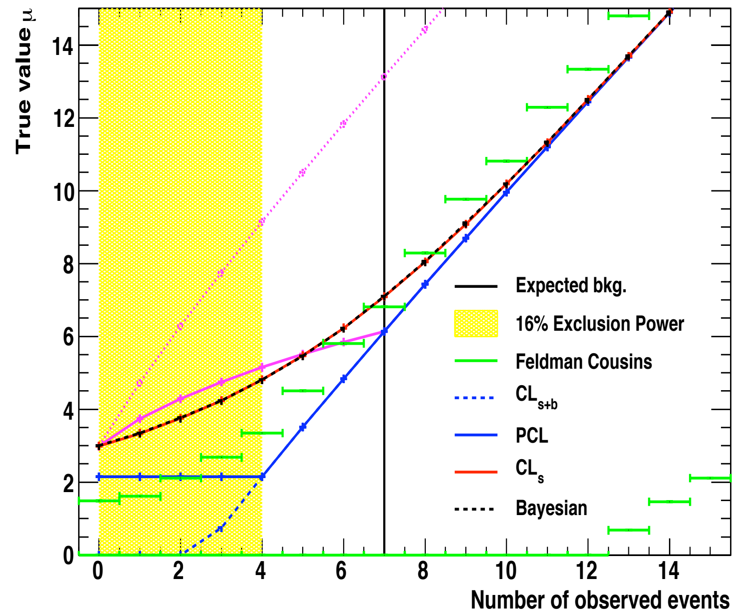
Confidence Intervals and Limits for Pedestrians

Markus Schumacher

GK Lecture, Freiburg, 26 - 28 September 2017





Goal of the lecture: understand the two figures

Outline

Lecture 1: Basics (26.9.)

- Motivation
- Frequentist and Bayesian Probability
- Parameter Estimation from Maximum Likelihood
- Frequentist Confidence Intervals a la Neyman and Coverage
- Bayesian Credibility Interval from Likelihood Principle

Lecture 2: Limits for Gaussian Probability Distribution (27.9)

- Connection of Frequentist Limit to Frequentist Hypothesis Test
- Limits close to physical boundary
- Frequentist and Bayesian Limits
- Modified Frequentist: CL_s Method and Power Constrained Limit (PCL)
- Unified Approach, Feldman- Cousins Intervals (FCL)

Lecture 3: Limits for Poisson Distribution (28.9.)

- Confidence Intervals
- Limits close to physical boundary
- Frequentist, Bayesian, PCL, CL_s, FC Limits

Outline

Lecture 1: Basics (26.9.)

- Motivation
- Frequentist and Bayesian Probability
- Parameter Estimation from Maximum Likelihood
- Frequentist Confidence Intervals a la Neyman and Coverage
- Bayesian Credibility Interval from Likelihood Principle

Lecture 2: Limits for Gaussian Probability Distribution (27.9)

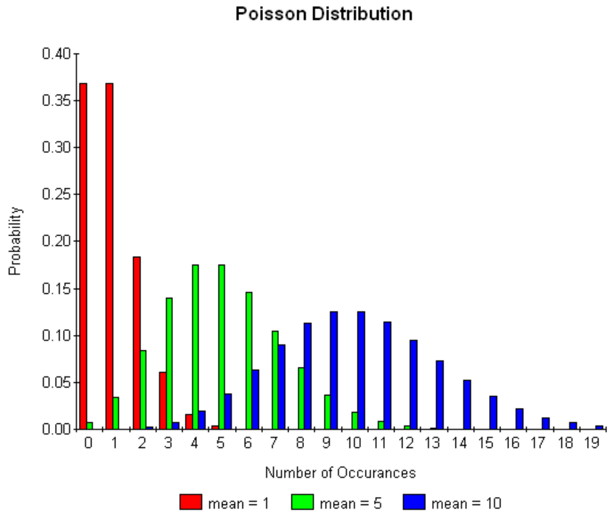
- Connection of Frequentist Limit to Frequentist Hypothesis Test
- Limits close to physical boundary
- Frequentist and Bayesian Limits
- Modified Frequentist: CL_s Method and Power Constrained Limit (PCL)
- Unified Approach, Feldman- Cousins Intervals (FCL)

Lecture 3 +4 : Limits for Poisson Distribution (28.+29.9.)

- Confidence Intervals
- Limits close to physical boundary
- Frequentist, Bayesian, PCL, CL_s, FC Limits

Confidence Intervals for Poisson-PDF

$$f(n;\lambda) = \frac{\lambda^n}{n!}\exp(-\lambda)$$



n = observed events = ML estimate for λ
Target: confidence interval for λ

Due to the discreteness of n the „confidence belt" equations can not be fulfilled exactly.
„Conservative" modification of equations e.g:

$$\alpha \geq P(\hat{\lambda} \geq u_\alpha(\lambda))$$

Hence over-coverage per construction

$$P(a \leq \lambda \leq b) \geq 1 - \alpha - \beta$$

Inversion of test

$$\alpha = \sum_{n=n_{obs}}^{\infty} f(n;a) = 1 - \sum_{n=0}^{n_{obs}-1} f(n;a) = 1 - \sum_{n=0}^{n_{obs}-1} \frac{a^n}{n!}e^{-a},$$

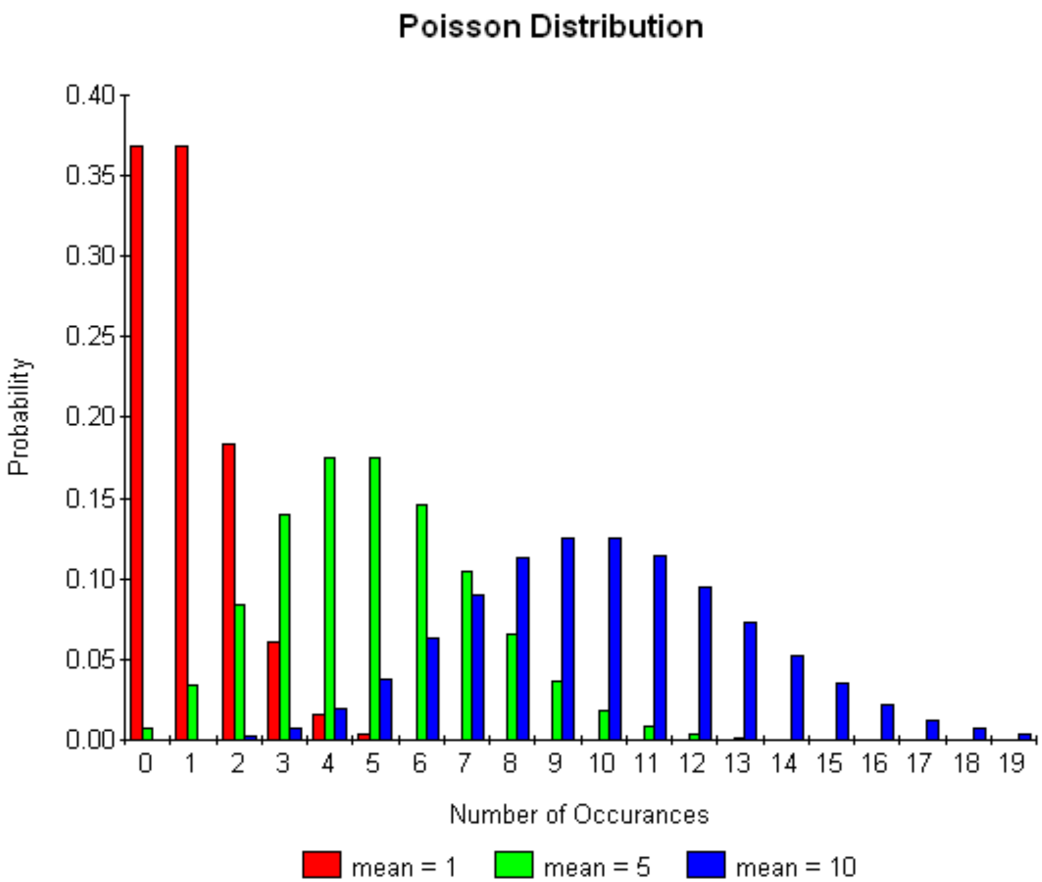
Solve numerically the equtions →

$$\beta = \sum_{n=0}^{n_{obs}} f(n;b) = \sum_{n=0}^{n_{obs}} \frac{b^n}{n!}e^{-b}.$$

Poisson PDF

$$f(n;\lambda) = \frac{\lambda^n}{n!}\exp(-\lambda)$$

n = nr. of observed events = ML estimate for λ
Target: confidence interval for λ



Determination of CI for Poisson-Parameter

Simple case: no observed event $\beta = e^{-b} \Longrightarrow b = -\log\beta$

hence at CL = 95% $b = -\log(0.05) = 2.996 \approx 3.$

For general case use relation btw. Poisson-PDF and Chi²-PDF

$$\sum_{n=0}^{n_{obs}} \frac{\lambda^n}{n!} e^{-\lambda} = \int_{2\lambda}^{\infty} f_{\chi^2}(z; n_{dof} = 2(n_{obs}+1))dz$$
$$= 1 - F_{\chi^2}(2\lambda; n_{dof} = 2(n_{obs}+1)),$$

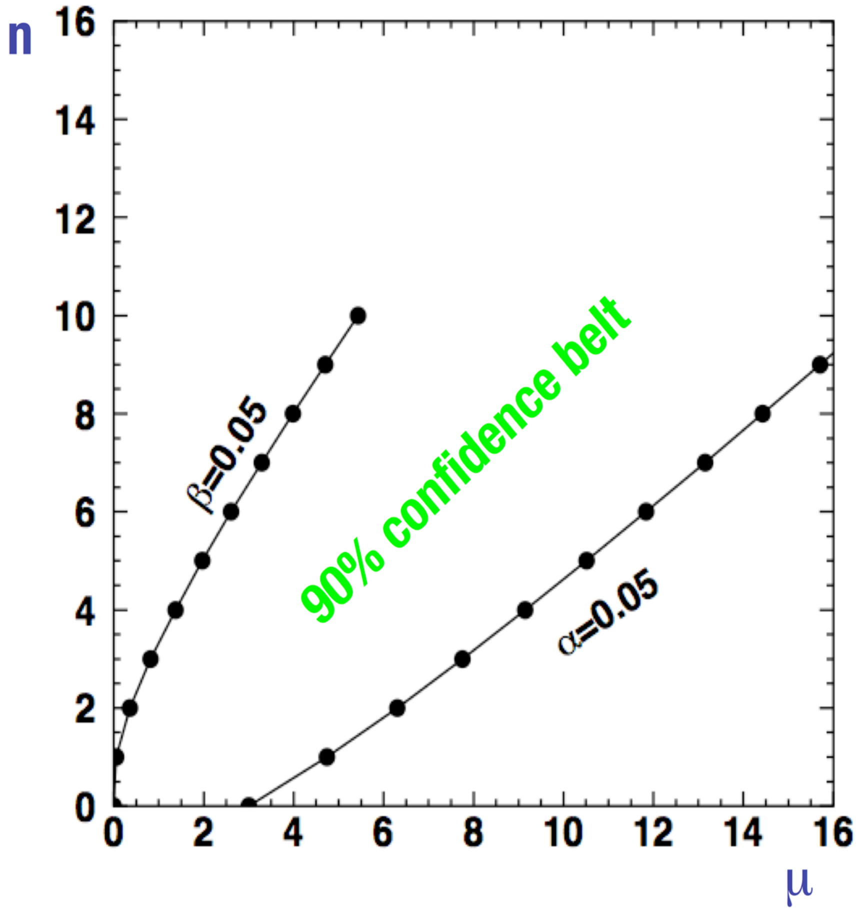
The borders of the CI are obtained via the cumulative of the Chi²-PDF

$$a = \frac{1}{2} F^{-1}_{\chi^2}(\alpha; n_{dof} = 2n_{obs}),$$
$$b = \frac{1}{2} F^{-1}_{\chi^2}(1-\beta; n_{dof} = 2(n_{obs}+1)).$$

n_{obs}	lower limit a			upper limit b		
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\beta = 0.1$	$\beta = 0.05$	$\beta = 0.01$
0	–	–	–	2.30	3.00	4.61
1	0.105	0.051	0.010	3.89	4.74	6.64
2	0.532	0.355	0.149	5.32	6.30	8.41
3	1.10	0.818	0.436	6.68	7.75	10.04
4	1.74	1.37	0.823	7.99	9.15	11.60
5	2.43	1.97	1.28	9.27	10.51	13.11
6	3.15	2.61	1.79	10.53	11.84	14.57
7	3.89	3.29	2.33	11.77	13.15	16.00
8	4.66	3.98	2.91	12.99	14.43	17.40
9	5.43	4.70	3.51	14.21	15.71	18.78
10	6.22	5.43	4.13	15.41	16.96	20.14

Frequentist Confidence Intervals for Poisson-PDF

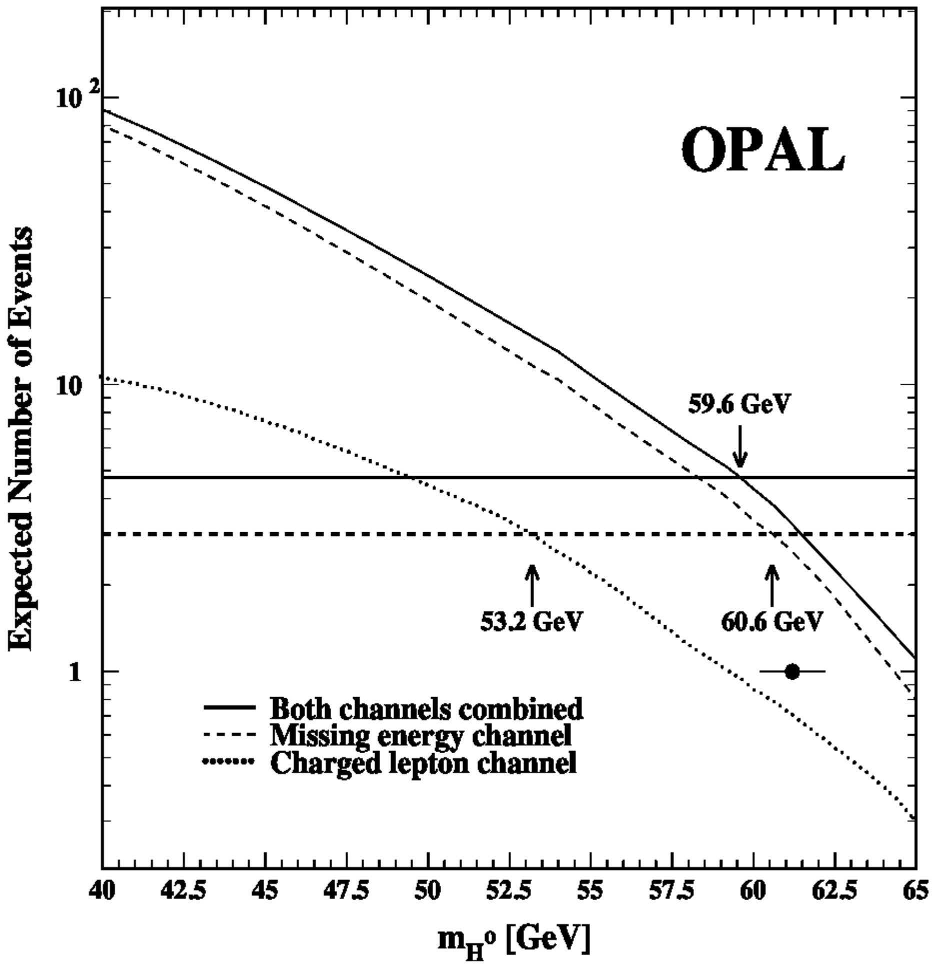
Confidence belt



Upper a and lower b limits

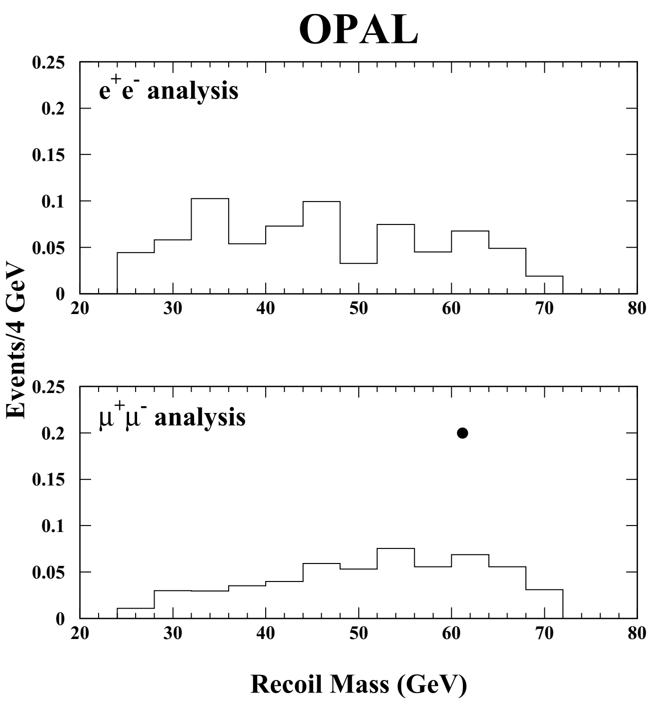
n_{obs}	lower limit a			upper limit b		
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\beta = 0.1$	$\beta = 0.05$	$\beta = 0.01$
0	-	-	-	2.30	3.00	4.61
1	0.105	0.051	0.010	3.89	4.74	6.64
2	0.532	0.355	0.149	5.32	6.30	8.41
3	1.10	0.818	0.436	6.68	7.75	10.04
4	1.74	1.37	0.823	7.99	9.15	11.60
5	2.43	1.97	1.28	9.27	10.51	13.11
6	3.15	2.61	1.79	10.53	11.84	14.57
7	3.89	3.29	2.33	11.77	13.15	16.00
8	4.66	3.98	2.91	12.99	14.43	17.40
9	5.43	4.70	3.51	14.21	15.71	18.78
10	6.22	5.43	4.13	15.41	16.96	20.14

Application in OPAL LEP 1 Higgs Boson Searches



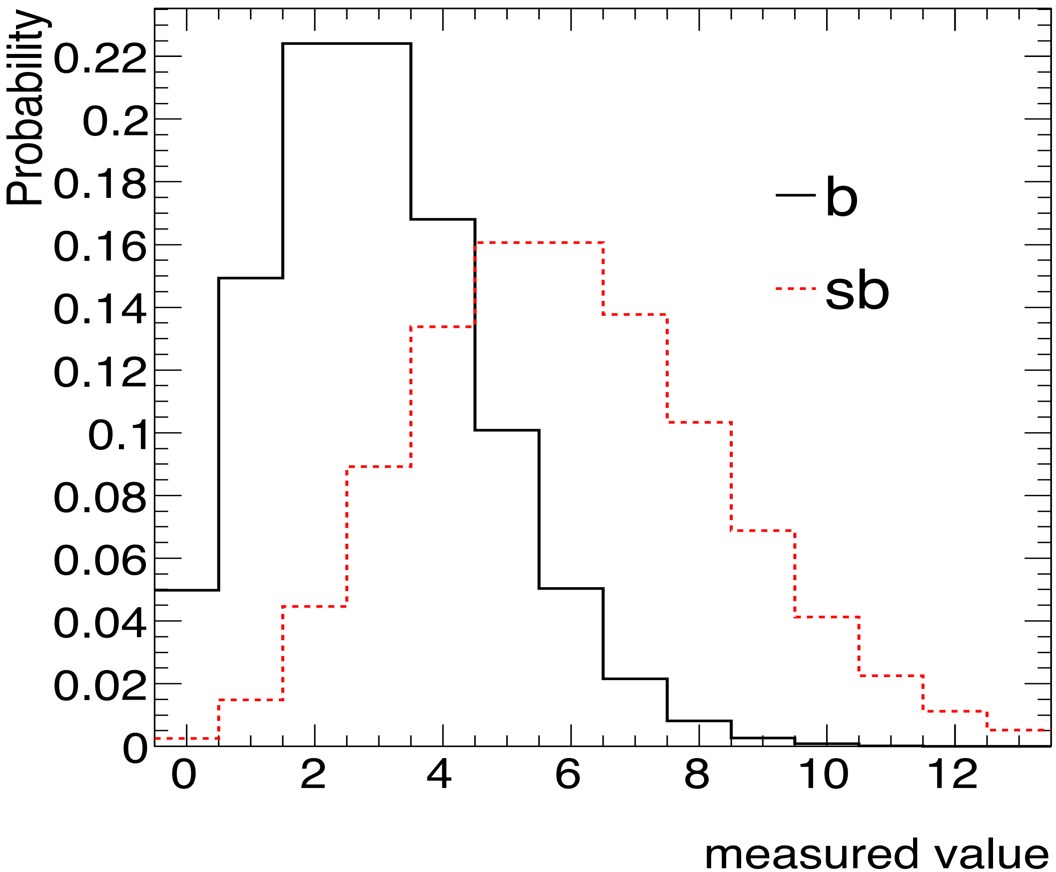
$n_{obs} = 0 \rightarrow s \leq 3.0$ at 95 %CL

$n_{obs} = 1 \rightarrow s \leq 4.7$ at 95% CL



$M_H < 59.6$ GeV excluded at 95% CL

Poisson-PDF for Signal plus Background



$$\boxed{P(n; s, b) = \frac{(s+b)^n}{n!} e^{-(s+b)}}$$

Expected known background rate b
Expected signal rate s to be estimated from data

Upper limit for Poisson-PDF with Background

Upper limit s at CL=1-γ given by solving the equation from test inversion

$$\gamma = P(n \leq n_{\text{obs}}; s, b) = \sum_{n=0}^{n_{\text{obs}}} \frac{(s+b)^n}{n!} e^{-(s+b)}$$

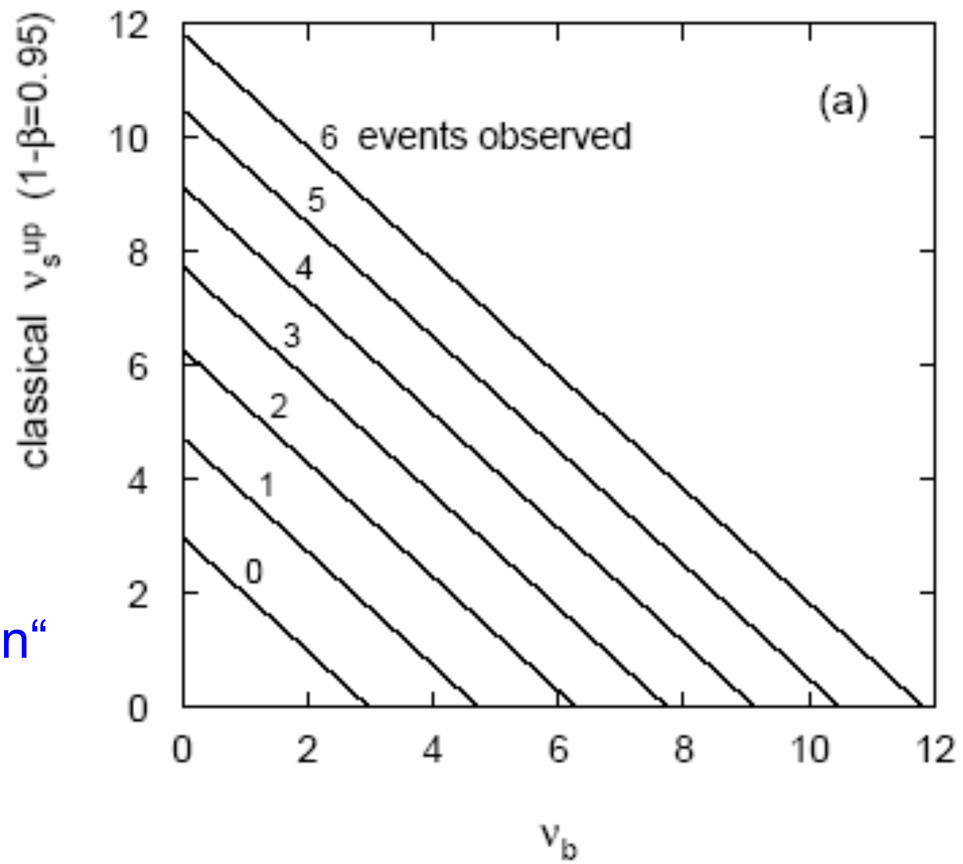
Boundaries of CI s_{lo}, s_{up} determined using Chi^2-PDF:

$$\boxed{\begin{aligned} s_{\text{lo}} &= \frac{1}{2} F^{-1}_{\chi^2}(\alpha; 2n) - b \\ s_{\text{up}} &= \frac{1}{2} F^{-1}_{\chi^2}(1-\beta; 2(n+1)) - b \end{aligned}}$$

same as for „b=0“ – b

→ called „background subtraction“

$n \leq b$ can yield $s_{up} < 0$



Frequentist Limit at Physical Boundary

e.g. for b = 2.5 and n_{obs} = 0 we find upper limit of $s_{\rm up} = -0.197 \quad (\mathrm{CL} = 0.90)$

increase CL to 0.95 yields s_{up} = 0.496

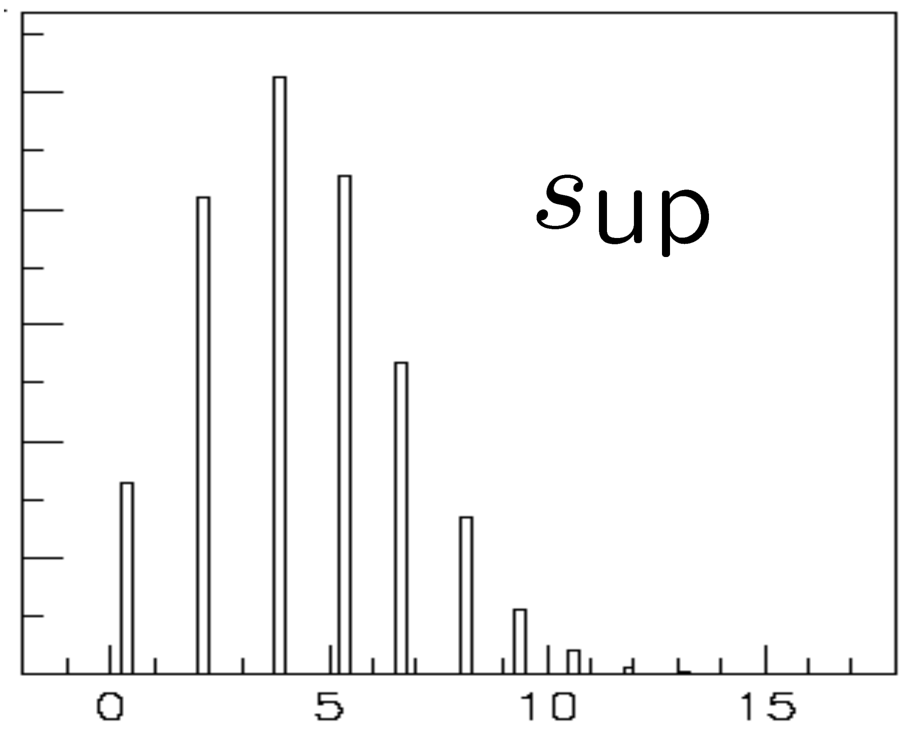
„cheating“ with CL = 0.917923 yields $s_{up} = 10^{-4}$!

naive argument: for b = 2.5 → variance is $\sqrt{2.5}$ = 1.6. how can limit be so small?

MC simulation:
determine median limit under „b-only“ hypothesis (s = 0)
→ expected limit

distribution of 95% CL upper limits for b = 2.5, s = 0.

→ Median s_{up} = 4.44



Bayesian Upper Limit for Poisson-PDF

Bayesian upper limit to CL = 1-α to be derived from

.

$$1-\alpha = \int_{-\infty}^{s_{\text{up}}} p(s|n)ds = \frac{\int_{-\infty}^{s_{\text{up}}} L(n|s)\,\pi(s)\,ds}{\int_{-\infty}^{\infty} L(n|s)\,\pi(s)\,ds}$$

with likelihood function

$$L(n|s) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

and uniform prior in physical region

$$\pi(s) = \begin{cases} 1 & s \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Posterior probability:

$$p(s|n) = \frac{(s+b)^n e^{-(s+b)}}{\Gamma(b,n+1)}$$

$$\Gamma(b,n+1) = \int_b^{\infty} x^n e^{-x}\,dx$$

Need so solve:

$$1-\alpha = \int_0^{s_{\text{up}}} p(s|n)\,ds$$

$$\alpha = e^{-s_{\text{up}}} \frac{\sum_{m=0}^{n} (s_{\text{up}}+b)^m/m!}{\sum_{m=0}^{n} b^m/m!}$$

Upper limit given by

$$s_{\text{up}} = \frac{1}{2} F_{\chi^2}^{-1}\left[p, 2(n+1)\right] - b$$

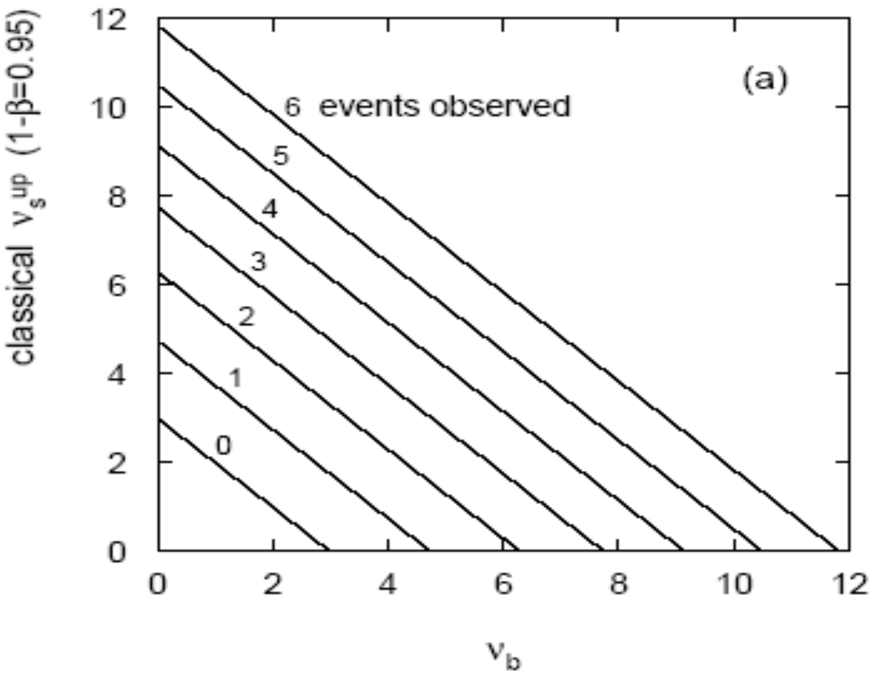
$$\int_0^a x^n e^{-x}\,dx = \Gamma(n+1) F_{\chi^2}(2a, 2(n+1))$$

Frequentist formula modified by replacing (1-α) by p

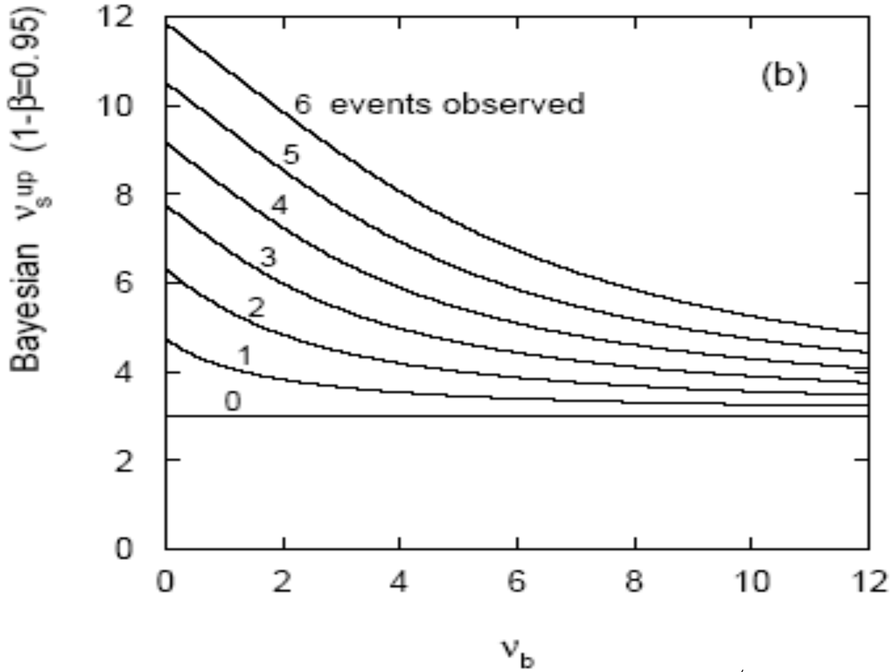
$$p = 1-\alpha\left(1 - F_{\chi^2}\left[2b, 2(n+1)\right]\right)$$

Classical and Bayesian Limits at 95% CL

Frequentist Limit at CL=95%



Bayesian Limit at CL=95%
with uniform prior



Upper limit can be „0“

$$\boxed{s_{up} = \frac{1}{2} F_{\chi^2}^{-1}(1-\beta; 2(n+1)) - b}$$

Upper limit always ≥ 3 $p = 1 - \alpha\left(1 - F_{\chi^2}[2b, 2(n+1)]\right)$

$$\boxed{s_{up} = \frac{1}{2} F_{\chi^2}^{-1}[p, 2(n+1)] - b}$$

for b= 0 identical

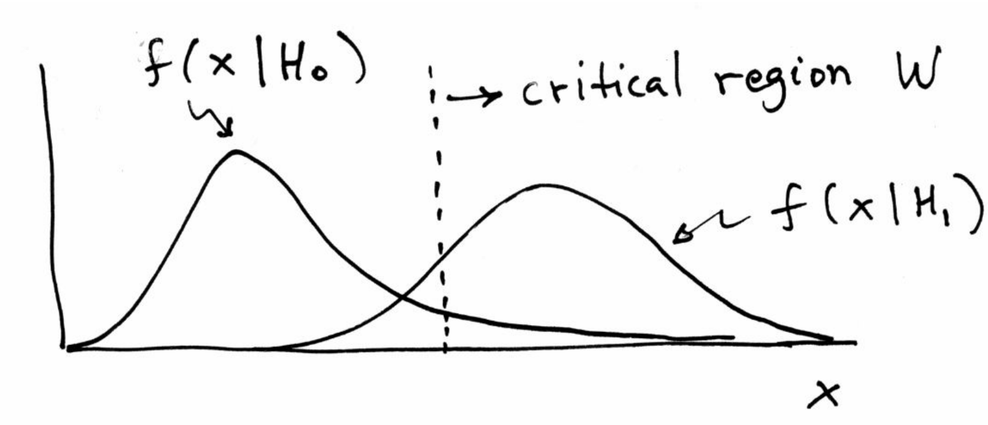
for n>>b also identical

other b values Bayesian> classical limit →“conservative” coverage > CL

Bayesian Limit independent on b for n= 0

Neyman Pearson Lemma

Best test: for given significance level α, maximize power M=1-β



Questions: Which test statistic t?
Which choice of critical region?

Simple hypothesis H_0 and H_1

Neyman-Person-Lemma: a test of a simple null hypothesis H_0 w.r.t. to the simple alternative hypothesis H_1 is a best test, if the critical region is chosen such that inside it holds:

$$t_{NP} = \frac{\mathcal{L}(x_{SP}|H_0)}{\mathcal{L}(x_{SP}|H_1)}$$

P = probability to observe sample x
(≤ c outside critical region)
c is a constant depending on α

Equivalent statement: the optimal test statistics is given by the likelihood ratio
(or any monotonic function 1/t(, t/(1+t), ln t)

$$t_{NP} < t_{crit}(\alpha)$$

Challenge in praxis: determination of PDFs for t under different hypothesis

Neyman Pearson Test Statistic for Exclusion

The Likelihood to observe n given H_0 (s=0,b) is:

$$L_b = \frac{b^n}{n!} e^{-b}$$

The Likelihood to observe n given H_1 (s,b) is:

$$L_{s+b} = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

→ Neyman-Pearson-Lemma: best test given by $\boxed{\frac{L_{s+b}}{L_b}}$

or monotonic function

$$\boxed{\ln \frac{L_{s+b}}{L_b} = n \ln\left(1 + \frac{s}{b}\right) - s}$$

Likelihood ratio is monotonic function of n.
PDF for optimal test statistic is also Poisson distribution

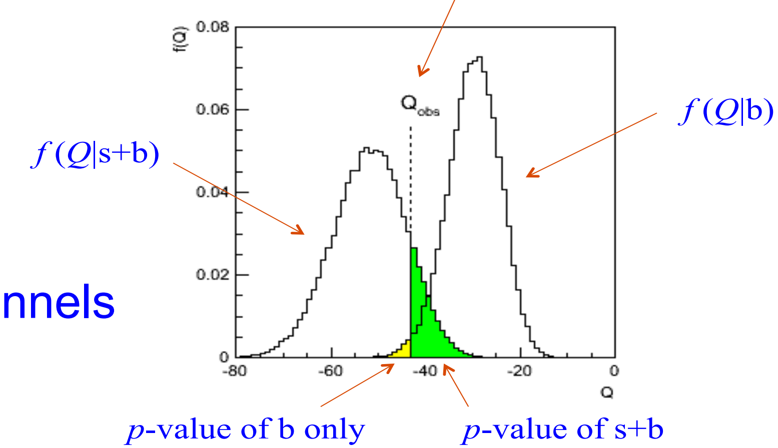
→ Counting rate n is optimal test statistic

Take e.g. b = 100, s = 20.

Suppose in real expe
Q is observed here.

Often used at LEP :

$$\boxed{Q = -2 \ln \frac{L_{s+b}}{L_b}}$$



Optimal use of distributions/ combination of channels
→ product of likelihoods per bin/channel
or sum of ln lik. per channel/bin

Profile Likelihood Test Statistic

Nullhypothesis H_0 simple, Alternative hypothesis H_1 composite

$$t_{PL} = \frac{\mathcal{L}(x_{SP}|H_0(\theta))}{\mathcal{L}(x_{SP}|H_1(\hat{\theta}_{ML}))}$$

- not mathematically proof that this the best, but in praxis no better found
- allows easy incorporating of syst. uncertainties via profiled nuisance parameters
- PDF for $q = -2 \ln t_{PL}$ is Chi2-PDF with 1 degree of freedom $f_{Chi2}(q; \nu=1)$

  for N_{SP} not too small (Wilks theorem)

two sided critical region / test recommended from application to particular problem

Profile Likelihood Test Statistic for Exclusion

So far: signal rate fixed (known) under alternative hypothesis
Now: find best number of signal events under H_1 via maximum likelihood fit
i.e. H_1 is composite hypothesis with signal count as free parameter

Likelihood function
$$L(n; s, b) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

Test statistic:
$$\boxed{\lambda(s) = \frac{L(s)}{L(\hat{s})}}$$

λ in [0, 1]:
1 good agreement with H_0

Enumerator (zähler): likelihood for H_0 (s fixed, for discovery s=0)
Denominator: likelihood for H_1 (s estimated from data)

Maximum likelihood estimate for signal counts: $\hat{s} = n - b$

Test statistics for discovery (s=0 in enumerator):

$$\boxed{\lambda(s) = n \ln(s+b) - (s+b) - n \ln n + n}$$

ln λ in [0, -infinity]:
0 good agreement with H_0

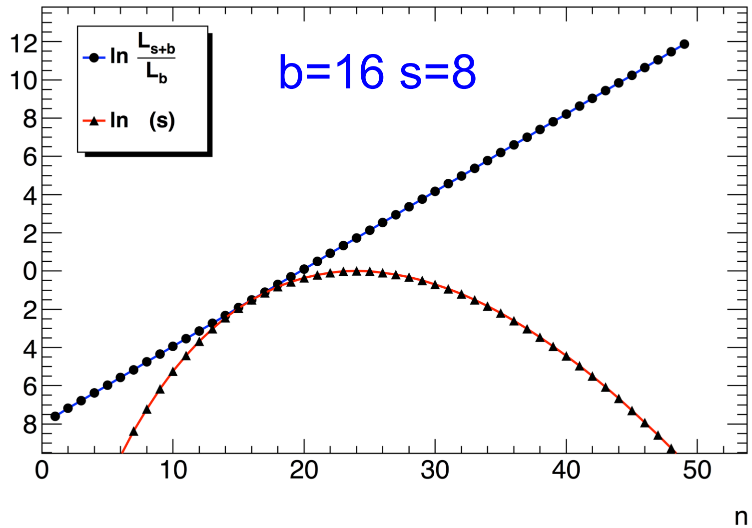
Comparison of Test Statistic for Exclusion

From Neyman-Pearson-Lemma
(simple hypothesis):

$$\ln \frac{L_{s+b}}{L_b} = n \ln\left(1 + \frac{s}{b}\right) - s$$

From profile likelihood ratio
(composite alternative hypothesis H_1)

$$\lambda(s) = n\ln(s+b) - (s+b) - n\ln n + n$$



If we consider a deviation from background only hypothesis only for n>b (e.g. set ln $\lambda(0)$ = 0 for n<b)

then both are monotonic and as optimal as using n
(for counting experiment neglecting systematic uncertainties)

ln λ(s) preferred for multiple channels / distributions
add values of ln λ (s) for each/bin channel

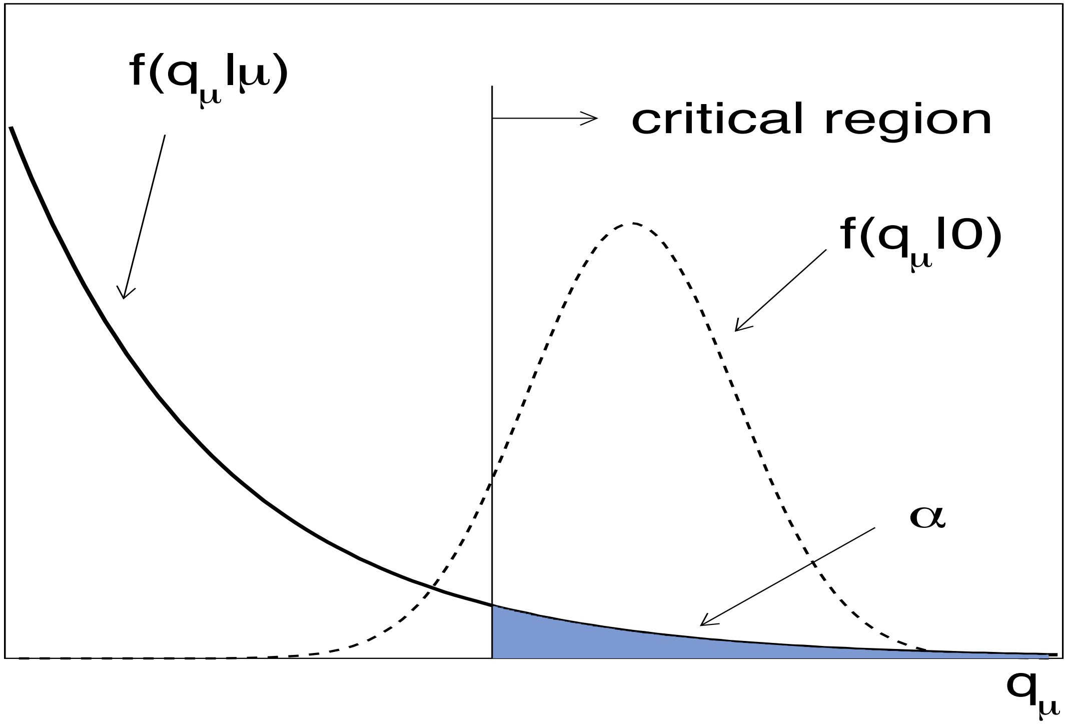
PDF for -2 ln λ(s) for „s+b“ /”b-only) given by Wilks’ (Wald’s)theorem

Profile Likelihood Test statistic for Exclusion

H_0: signal+background → μ=1 , b H_1: background only μ = 0 , b
μ parametrises strength w.r.t. "standard prediction" $\mu = s_{obs}/s_{SM}$

Test statistic $q_\mu = - 2 \ln (\mu)$

One sided test, only signal strength < μ considered as inconsistent with H_0



large values H_1-like

small values H_0-like

$\mu^{95\%\mathrm{CL}}$

decrease tested μ until
P-value = α = 1- CL

„Spurious Exclusion“ with Frequentist Limit

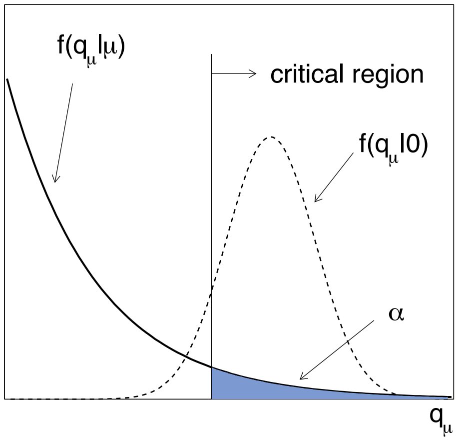
$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} \,|\, \text{signal+background}) = \int_{\tilde{q}_\mu^{\rm obs}}^{\infty} f(\tilde{q}_\mu|\mu, \hat{\theta}_\mu^{\rm obs})\, d\tilde{q}_\mu$$

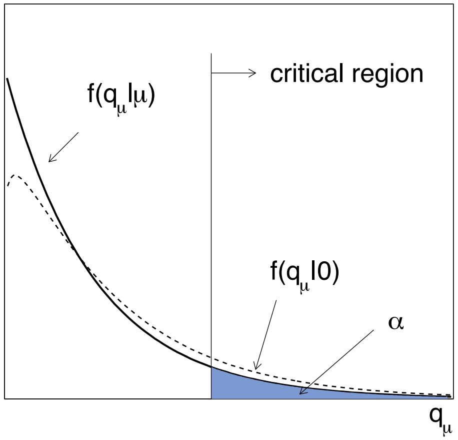
Pure frequentist would stop and say: „signal + background“ hypothesis is excluded with a confidence level CL_{S+B} of 1- p_μ

„Problem“: Spurious exclusion of signals with no sensitivity (s<<b)

large s

power M= 1-ß large w.r.t. significance level α





s<<b

powerM = 1-ß ~ significance level α

signal+BG-like ← → BG only like,

By construction: probability to reject μ if μ is true is α

for s<<b probability to reject very small μ if μ=0 is true ~ α + epsilon

→ probability to exclude hypotheses with zero signal (due to downwards fluctuation) ~ α „spurious exclusion w/o sensitivity“

Pseudo-Frequentist or Zech's Interpretation

Bayesian limit with uniform prior first proposed by O. Helene (1983)
Condition can be rewritten as

$$\boxed{\alpha = e^{-s_{\mathrm{up}}} \frac{\sum_{m=0}^{n} (s_{\mathrm{up}} + b)^m / m!}{\sum_{m=0}^{n} b^m / m!}}$$

Numerical identical result derived by G. Zech (1988) in different context

$$P(n; s+b) = \frac{e^{-(s+b)}(s+b)^n}{n!}$$

stems from

$$P(n; s+b) = \sum_{n_b=0}^{n} \sum_{n_s=0}^{n-n_b} P(n_b; b) P(n_s; s)$$

If N< b we know background in data < b
→ renormalilze background PDF
and replace it in compound PDF

$$P'(n_b; b) = P(n_b; b) / \sum_{n_b=0}^{N} P(n_b; b)$$

Find upper limit s by solving (with ε=α)

$$\boxed{\epsilon = \sum_{n=0}^{N} P(n; s+b) \Big/ \sum_{n_b=0}^{N} P(n_b; b)}$$

Zech's interpetation →
(not accepted by many Frequentist
as one conditions on data, but known
as the PDG formula for many years)

different. The limit in the "frequency interpretation" can be stated as follows: for an infinitely large number of experiments, looking for a signal with expectation s and Poisson distributed background with mean b, where the background is restricted to values of less than or equal to N, the frequency of observing N or less events is ϵ.

Zech's Interpretation

$$\epsilon = \sum_{n=0}^{N} P(n; s+b) \Big/ \sum_{n_b=0}^{N} P(n_b; b)$$

different. The limit in the "frequency interpretation" can be stated as follows: for an infinitely large number of experiments, looking for a signal with expectation s and Poisson distributed background with mean b, where the background is restricted to values of less than or equal to N, the frequency of observing N or less events is ϵ.

(not accepted by many Frequentist as one conditions on data,
but known as the PDG formula for many years)

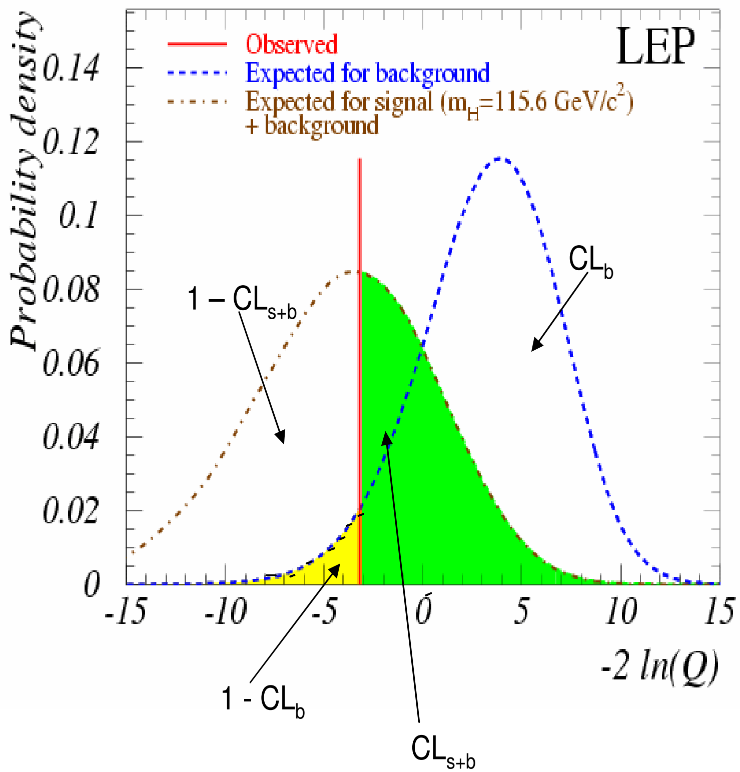
CL_S Limit for Poisson

A. Read (1997): applied Zech's "background conditioning" to the LEP test statistic Q
CL_S ≈ "confidence in the signal-only hypothesis"

$$CL_{s+b} = P_{s+b}(Q \le Q_{obs})$$

$$CL_b = P_b(Q \le Q_{obs})$$

$$CL_s \equiv CL_{s+b}/CL_b.$$



A hypothesis is exlcuded at confidence level CL if

$$1 - CL_s \le CL$$

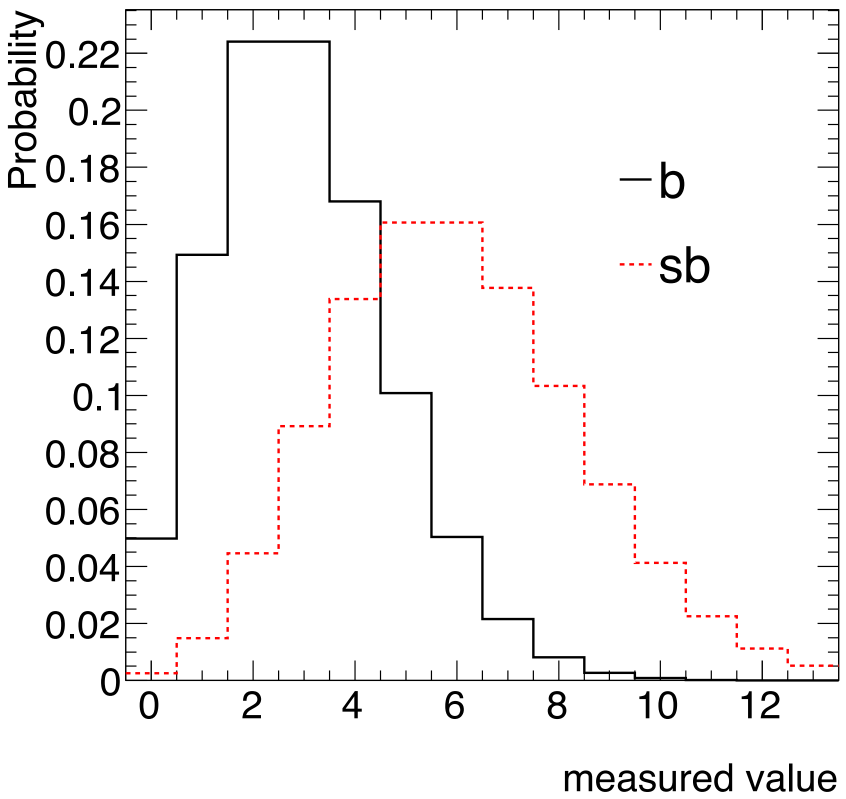
Applied to Poisson case yields Zech's formula:

$$CL_s = \frac{P(X \le X_{obs})}{P(X_b \le X_{obs})} = \frac{P(n \le n_{obs})}{P(n_b \le n_{obs})} \qquad CL = 1 - \frac{\sum_{n=0}^{n_{obs}} \frac{e^{-(b+s)}(b+s)^n}{n!}}{\sum_{n=0}^{n_{obs}} \frac{e^{-b}b^n}{n!}}.$$

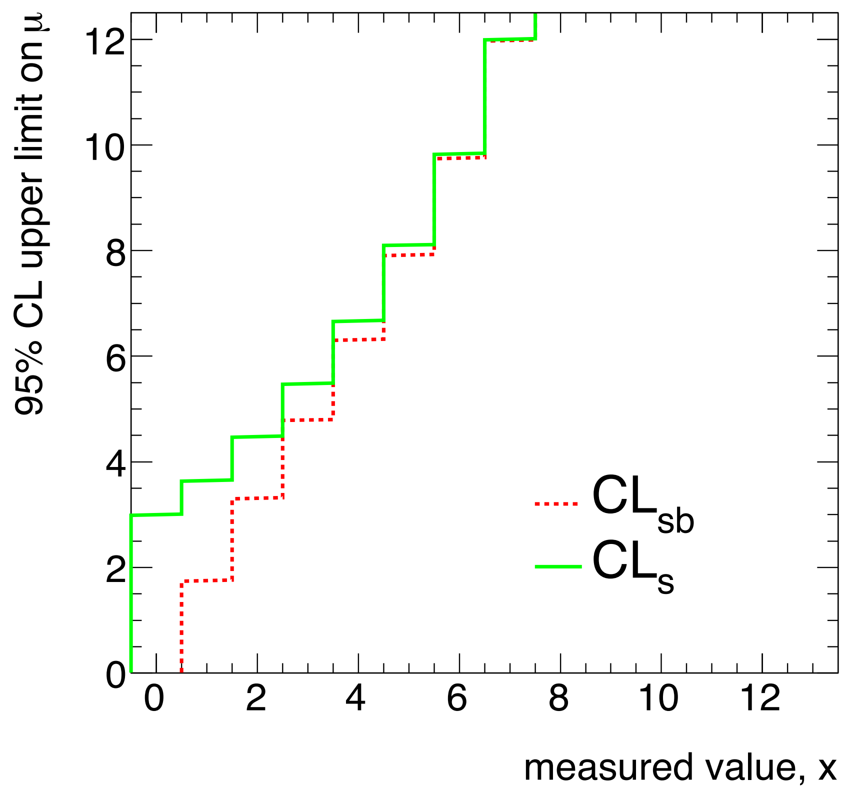
Remark: denominator is not 1-p-value for the b-only hyp. The sum would only run from 0 up to n_{obs}-1. Calling it the power is correct (I think)

Classical and CL_S Limit compared for Poisson PDF

Expected background b = 3
Expected signal yield s = 3



Upper limits from
classical approach CL_{sb}
CL_S technique



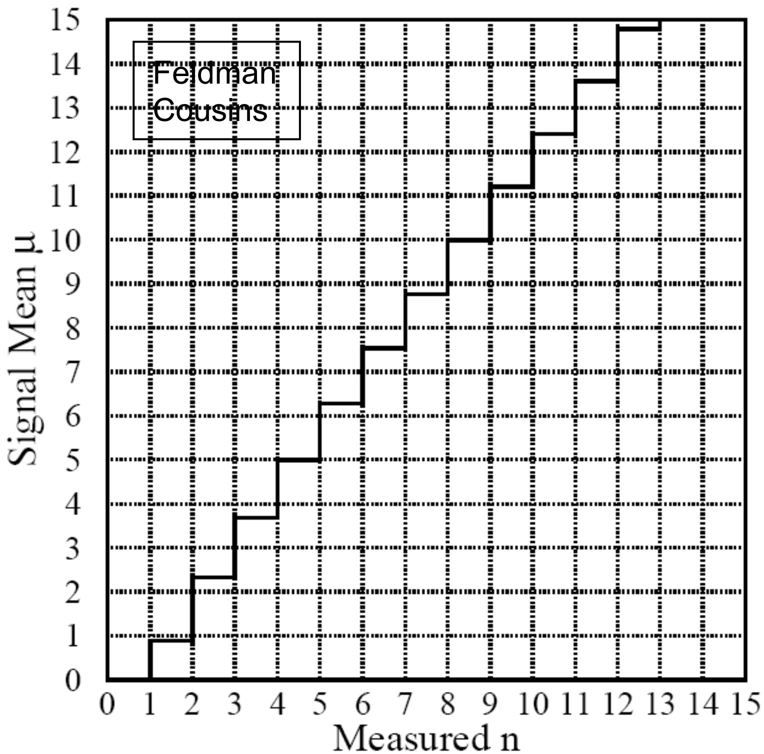
Flip-Flop-Problem for Poisson-Parameter s=μ

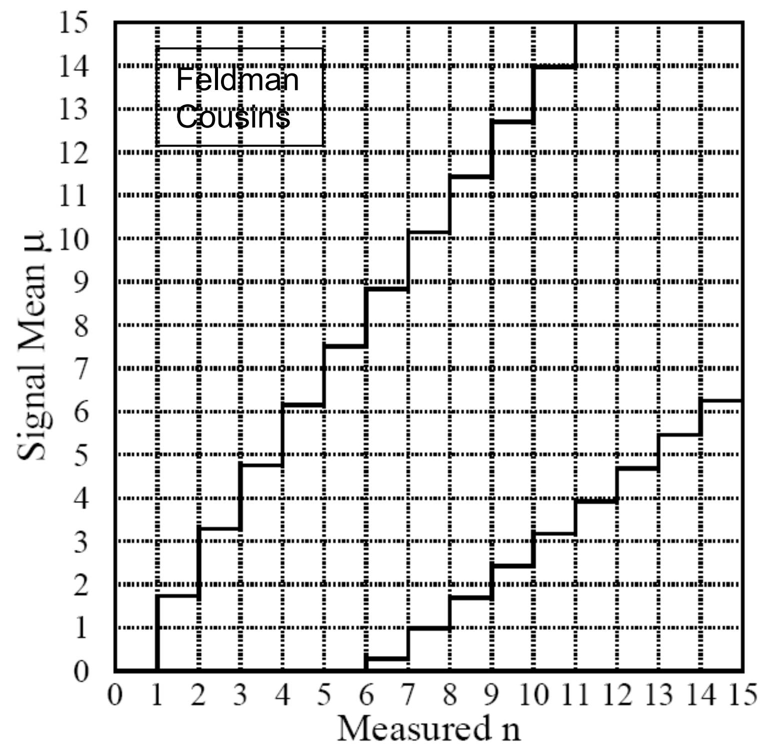
Known background =3

$$P(n|\mu) = (\mu+b)^n \exp(-(\mu+b))/n!$$

One-sided CI at CL=90%

Two-sided CI at CL=90%





For „Flip-Flop“ again to small coverage

Construction of confidence belt via likelihood ratio

$$l(s) = \frac{L(n|s,b)}{L(n|\hat{s},b)} \quad \text{where} \quad \hat{s} = \begin{cases} n-b & n \geq b, \\ 0 & \text{otherwise} \end{cases}$$

„Unified Approach": Poisson-CI at 90% CL

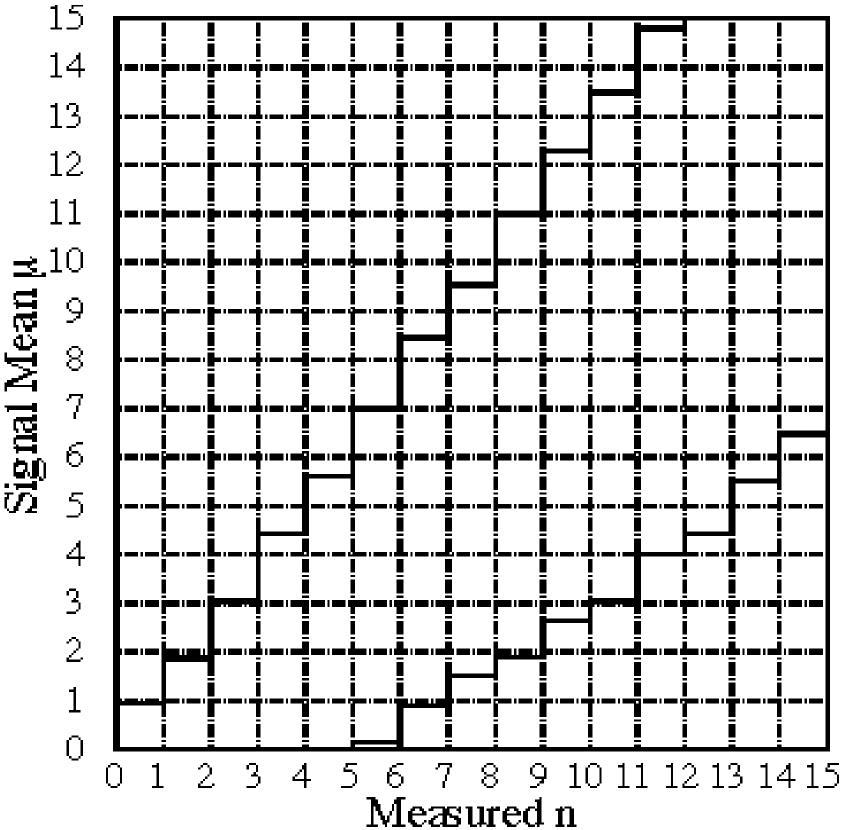
Construction of confidence belt
for μ=0.5, b=3

confidence belt for b=3

$$R = P(n|\mu)/P(n|\mu_{\text{best}})$$

Standard

x	$P(x\|\mu)$	$\hat{\mu}$	$P(x\|\hat{\mu})$	R	rank	U.L.	C.L.
0	0.030	0.0	0.050	0.607	6 •		
1	0.106	0.0	0.149	0.708	5 •	•	•
2	0.185	0.0	0.224	0.826	3 •	•	•
3	0.216	0.0	0.224	0.963	2 •	•	•
4	0.189	1.0	0.195	0.966	1 •	•	•
5	0.132	2.0	0.175	0.753	4 •	•	•
6	0.077	3.0	0.161	0.480	7 •	•	•
7	0.039	4.0	0.149	0.259		•	•
8	0.017	5.0	0.140	0.121		•	



„Unified Approach“: Poisson-CI at 90% CL

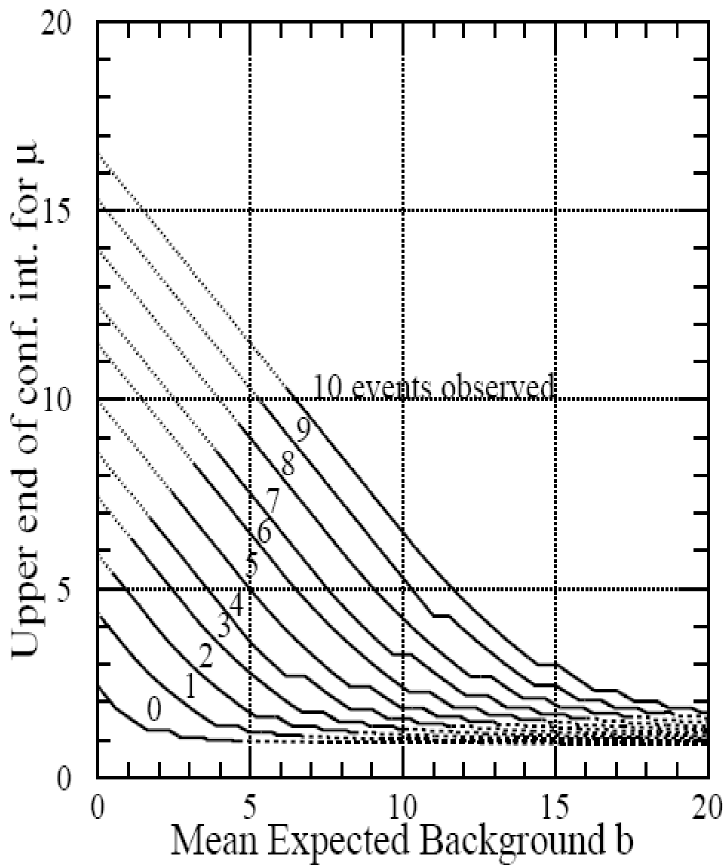


FIG. 8. Upper end μ_2 of our 90% C.L. confidence intervals $[\mu_1, \mu_2]$, for unknown Poisson sign mean μ in the presence of expected Poisson background with known mean b. The curves for th cases n_0 from 0 through 10 are plotted. Dotted portions on the upper left indicate regions whe μ_1 is non-zero (and shown in the following figure). Dashed portions in the lower right indica regions where the probability of obtaining the number of events observed or fewer is less than 1% even if $\mu = 0$.

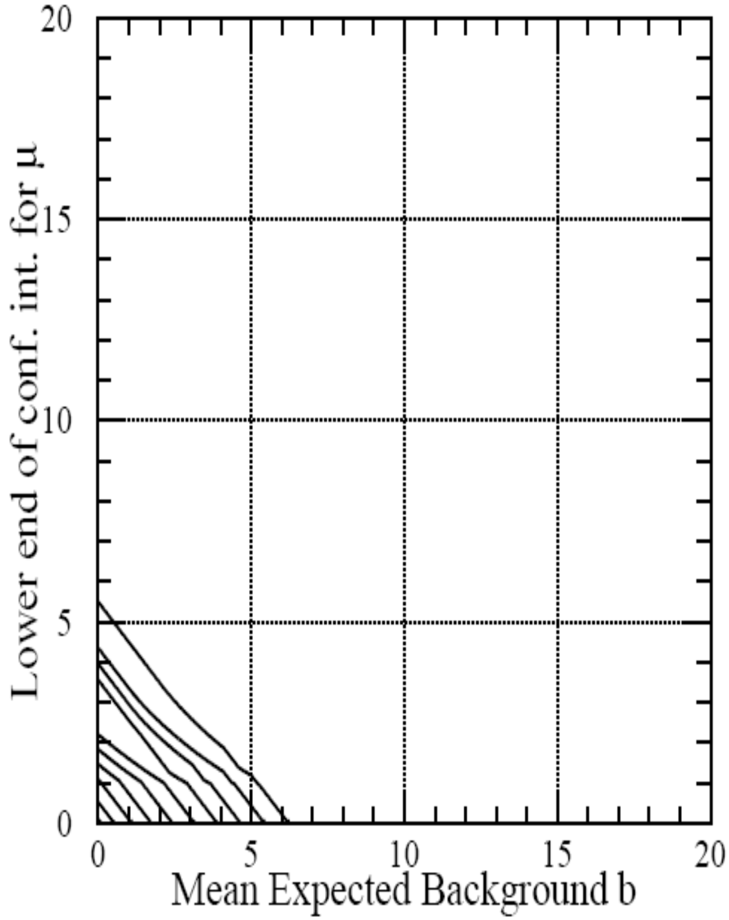


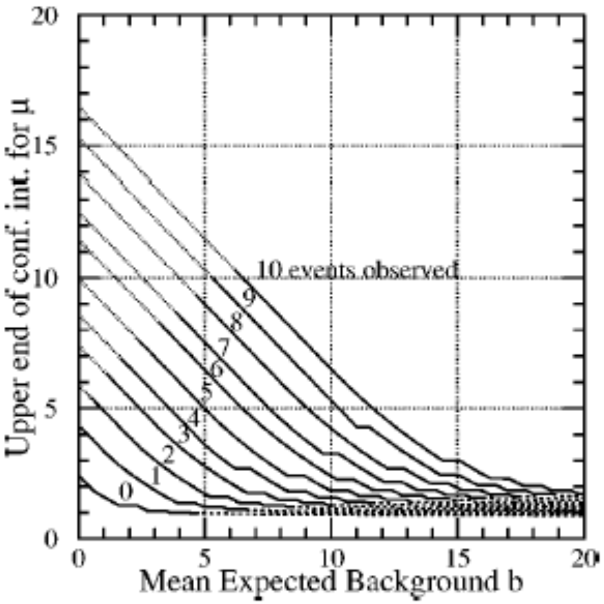
FIG. 9. Lower end μ_1 of our 90% C.L. confidence intervals $[\mu_1, \mu_2]$, for unknown Poisson signal mean μ in the presence of expected Poisson background with known mean b. The curves correspond to the dotted regions in the plots of μ_2 of the previous figure, with again $n_0 = 10$ for the upper right curve, etc.

Classical and Feldman Cousins Intervals

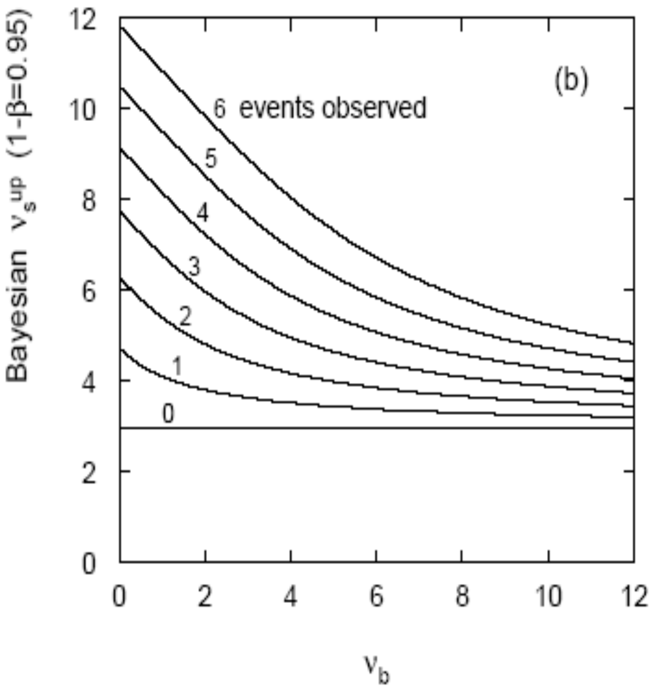
	Classic Frequentist		Feldman–Cousins	
	Upper limit	Equal-tailed interval		
N	95% CL	68% CL	95% CL	68% CL
0	3.00	[0.00, 1.84]	[0.00, 3.09]	[0.00, 1.29]
1	4.74	[0.17, 3.30]	[0.05, 5.14]	[0.37, 2.75]
2	6.30	[0.71, 4.64]	[0.36, 6.72]	[0.74, 4.25]
3	7.75	[1.37, 5.92]	[0.82, 8.25]	[1.10, 5.30]
4	9.15	[2.09, 7.16]	[1.37, 9.76]	[2.34, 6.78]
5	10.51	[2.84, 8.38]	[1.84, 11.26]	[2.75, 7.81]
6	11.84	[3.62, 9.58]	[2.21, 12.75]	[3.82, 9.28]
7	13.15	[4.42, 10.77]	[2.58, 13.81]	[4.25, 10.30]
8	14.43	[5.23, 11.95]	[2.94, 15.29]	[5.30, 11.32]
9	15.71	[6.06, 13.11]	[4.36, 16.77]	[6.33, 12.79]
10	16.96	[6.89, 14.27]	[4.75, 17.82]	[6.78, 13.81]

Comparison of Different Intervals

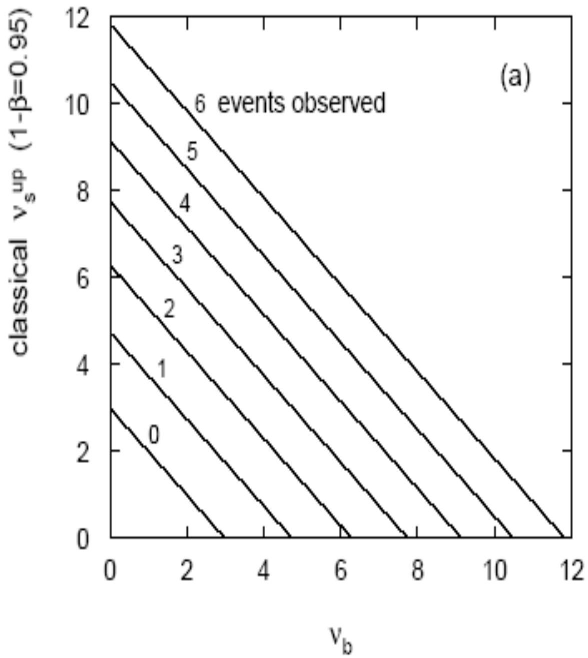
“Unified Approach”



Bayes/Zech/CLS

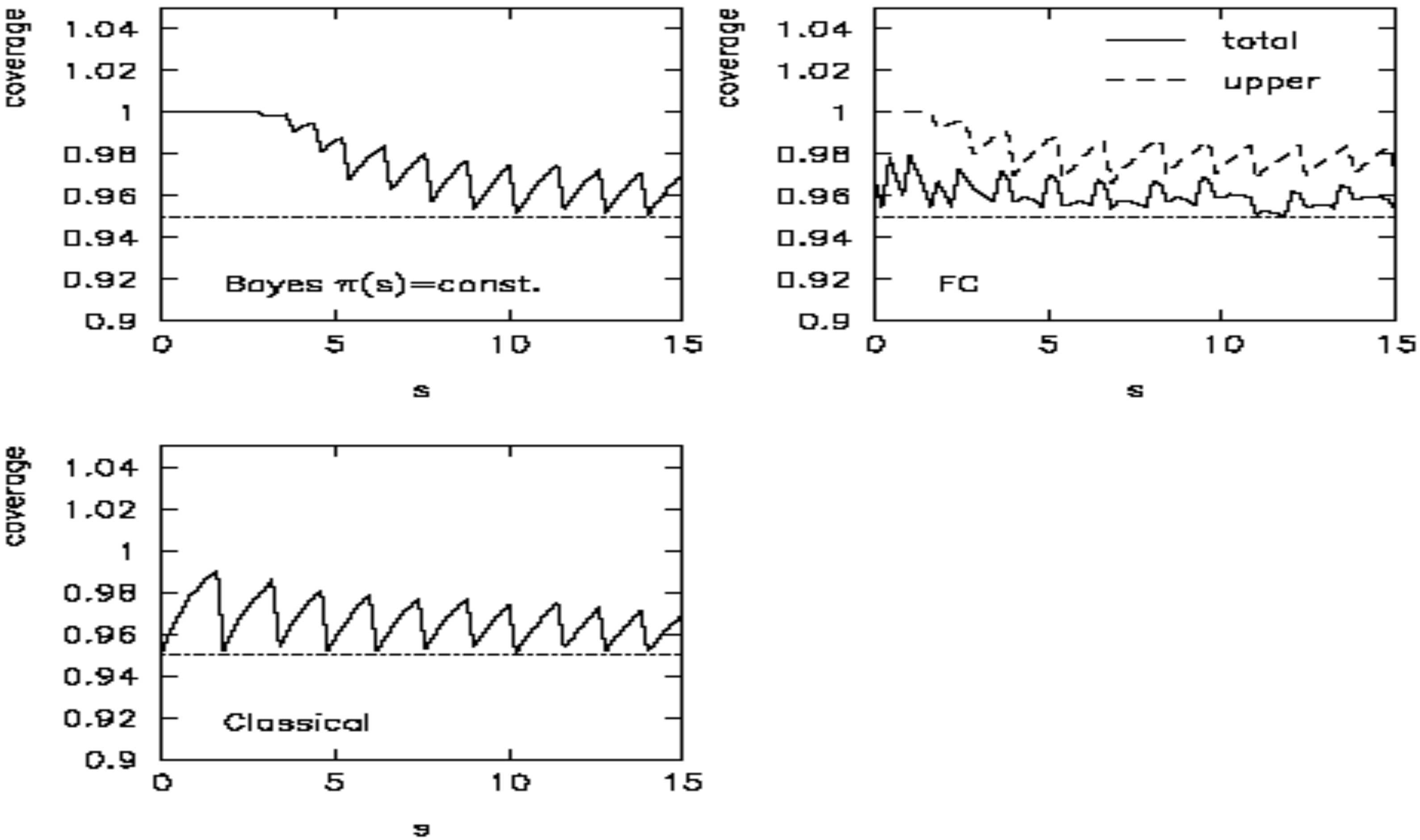


Classic Frequentist



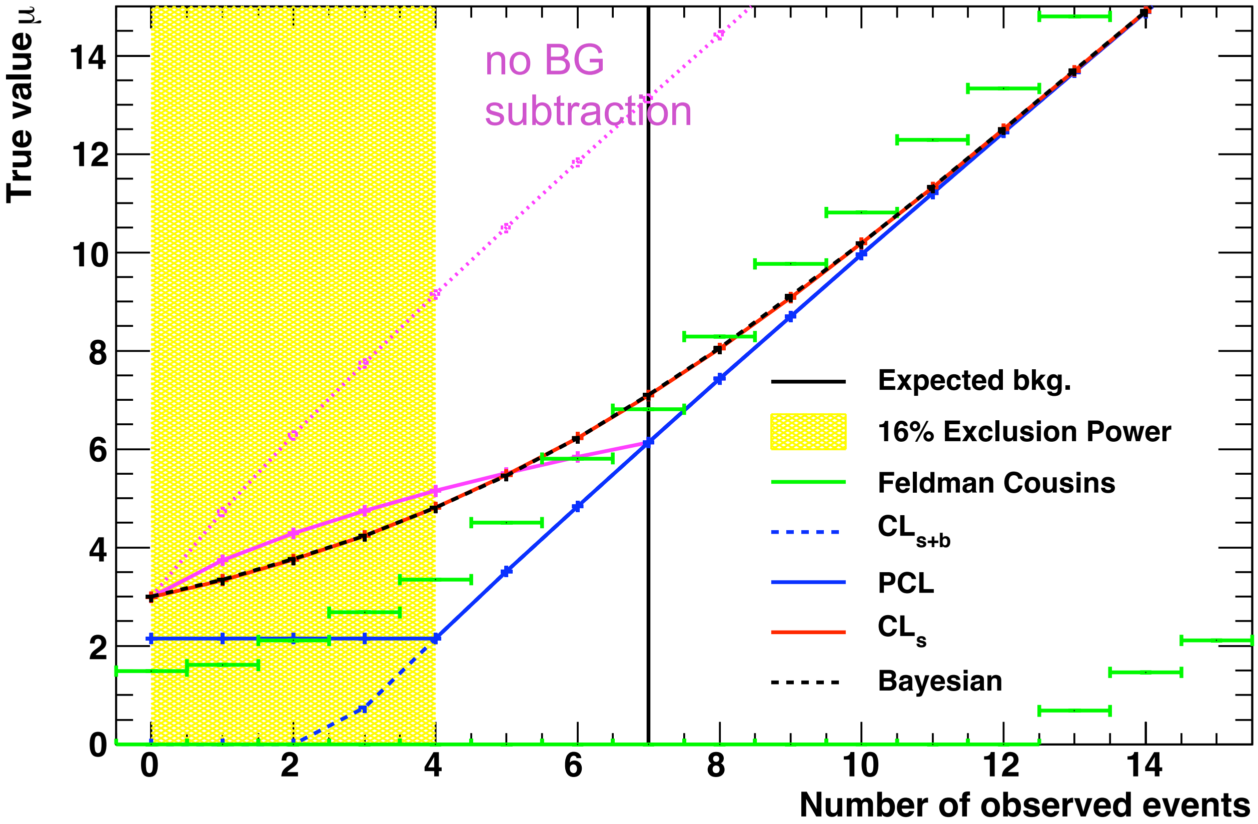
Coverage of Different Limits

Due to discrete nature of Poisson random variable the coverage is per construction larger than quoted CI also for Frequentist methods for most true values



Comparison of Different Limits for Poisson Mean

Simple counting experiment with exactly known background expectation of 7 events



- CL_S , Zech and Bayesian limit with flat prior in signal rate mathematically identical in praxis also very similar results for test statistics used at LHC (Tevatron, LEP)
- PCL= power constrained limit: require that power ≥ 16% (cut off at expected -1σ)

Conclusion of Lecture Series

In Limit of large event Samples and not close to a physical boundary

- Frequentist CI and Bayesian CI from flat prior agree numerically
- but the interpetation is always different
- CI = estimate ±1 standard deviation is a good aproximation for CI at 68%

Frequentist Limits

- Coverage Probabilty of quoted CL is guiding principle
- Neyman construction of confidence belt is cumbersome
- For many cases CI can be obtained from inversion of hypothesis test
- „Empty CI“ not a problem in principle
- „Empty“ CI can be avoided by PCL, CL_S and FC limits
- Ad-hoc correction of PCL, CL_S „punish“ outcomes with small power for discrimination between μ_{up} and μ_0, but violate the coverage interpretation
- Unified approach with FC limits circumvent the „flip flop problem“

Bayesian limits

- Simple calculation based on integration of posterior probabilty
- Likelihood principle is the main focus. Coverage in principle not interesting
- Choice of prior is as always a matter of taste and debate (flat, Jeffrey‘s,..)
- Numerically identical to CL_S limits for Poisson and Gauss PDF

Conclusion of Lecture Series

In Limit of large event Samples and not close to a physical boundary

- Frequentist CI and Bayesian CI from flat prior agree numerically
- but the interpetation is always different
- CI = estimate ±1 standard deviation is a good aproximation for CI at 68%

Frequentist Limits

- Coverage Probabilty of quoted CL is guiding principle
- Neyman construction of confidence belt is cumbersome
- For many cases CI can be obtained from inversion of hypothesis test
- „Empty CI“ not a problem in principle
- „Empty“ CI can be avoided by PCL, CL_S and FC limits
- Ad-hoc correction of PCL, CL_S „punish“ outcomes with small power for discrimination between μ_{up} and μ_0, but violate the coverage interpretation
- Unified approach with FC limits circumvent the „flip flop problem“

Bayesian limits

- Simple calculation based on integration of posterior probabilty
- Likelihood principle is the main focus. Coverage in principle not interesting
- Choice of prior is as always a matter of taste and debate (flat, Jeffrey‘s,..)
- Numerically identical to CL_S limits for Poisson and Gauss PDF

Final Words

Dr. John Watson: I wonder what desperate circumstances could occasion such an appeal.

Sherlock Holmes: I have devised seven separate explanations, each of which would cover the facts as far as we know them.

Dr. John Watson: Oh, and which one do you favour, Holmes?

Sherlock Holmes: At the moment, I have no favourites.
Data, data, data! I cannot make bricks without clay!

Dr. John Watson: We cannot theorize without data, I'm afraid.

(A. C. Doyle, The Copper Beeches)