Minimum Power for PCL

ATLAS Statistics Forum
EVO, 10 June, 2011

Glen Cowan*
Physics Department
Royal Holloway, University of London
www.pp.rhul.ac.uk/~cowan
g.cowan@rhul.ac.uk

* with Kyle Cranmer, Eilam Gross, Ofer Vitells

Outline

Quick review of Power-Constrained Limits

For more details see: G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Power-Constrained Limits*, arXiv:1105.3166

Choice of minimum power threshold:

rationale for 16%

rationale for 50%

Review of Power-Constrained Limits

The basic idea is to consider a parameter value μ as excluded if two criteria are satisfied:

(a) The value μ is rejected by a statistical test, e.g., the *p*-value of μ is found less than $a = 5\%$.

The highest value not excluded is the *unconstrained* limit, μ_{up}.

(b) The sensitivity to μ exceeds a specified threshold.

The measure of sensitivity is the power $M_0(\mu)$ of the test of μ with respect to the alternative $\mu = 0$, i.e., the probability to reject μ if $\mu = 0$ is true.

I.e. require $M_0(\mu) \geq M_{min}$ for some appropriately chosen power threshold M_{min}.

The highest value not excluded by both (a) and (b) is the power-constrained limit, μ^*_{up}.

Illustration of low/high sensitivity

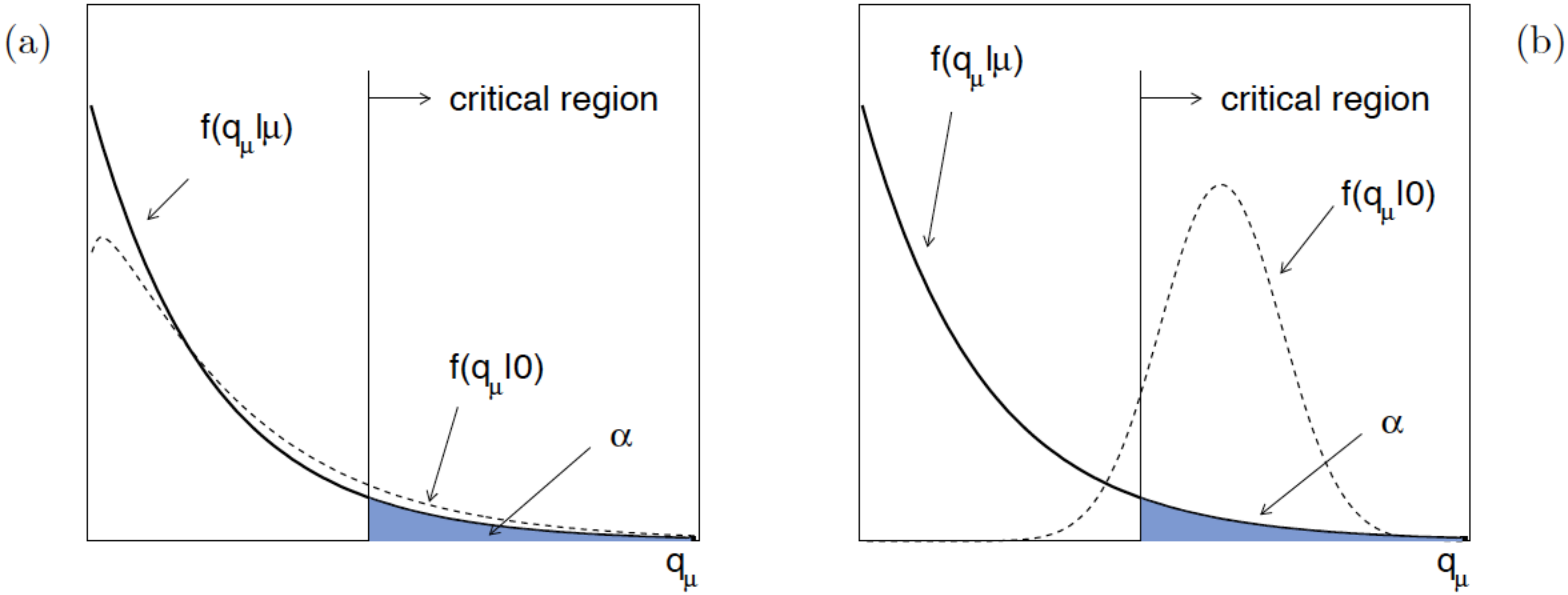


Figure 1: Illustration of statistical tests of parameter values μ for the cases of (a) little sensitivity and (b) substantial sensitivity (see text).

Power $M_0(\mu)$ is the area under the dashed curve in the critical region.

Choice of minimum power

Choice of M_{min} is convention; earlier we have proposed

$$M_{\text{min}} = \Phi(-1) = 0.1587$$

because in Gaussian example this means that one applies the power constraint if the observed limit fluctuates down by one standard deviation.

In fact the distribution of μ_{up} is often roughly Gaussian, so we call this a "1σ" (downward) fluctuation and use $M_{\text{min}} = 0.16$ regardless of the exact distribution of μ_{up}.

For the Gaussian example, this gives $\mu_{\text{min}} = 0.64\sigma$.

This is reasonable – the lowest limit is of the same order as the intrinsic resolution of the measurement (σ).

PCL and CLs/Bayesian for $x \sim$ Gauss(μ, σ)

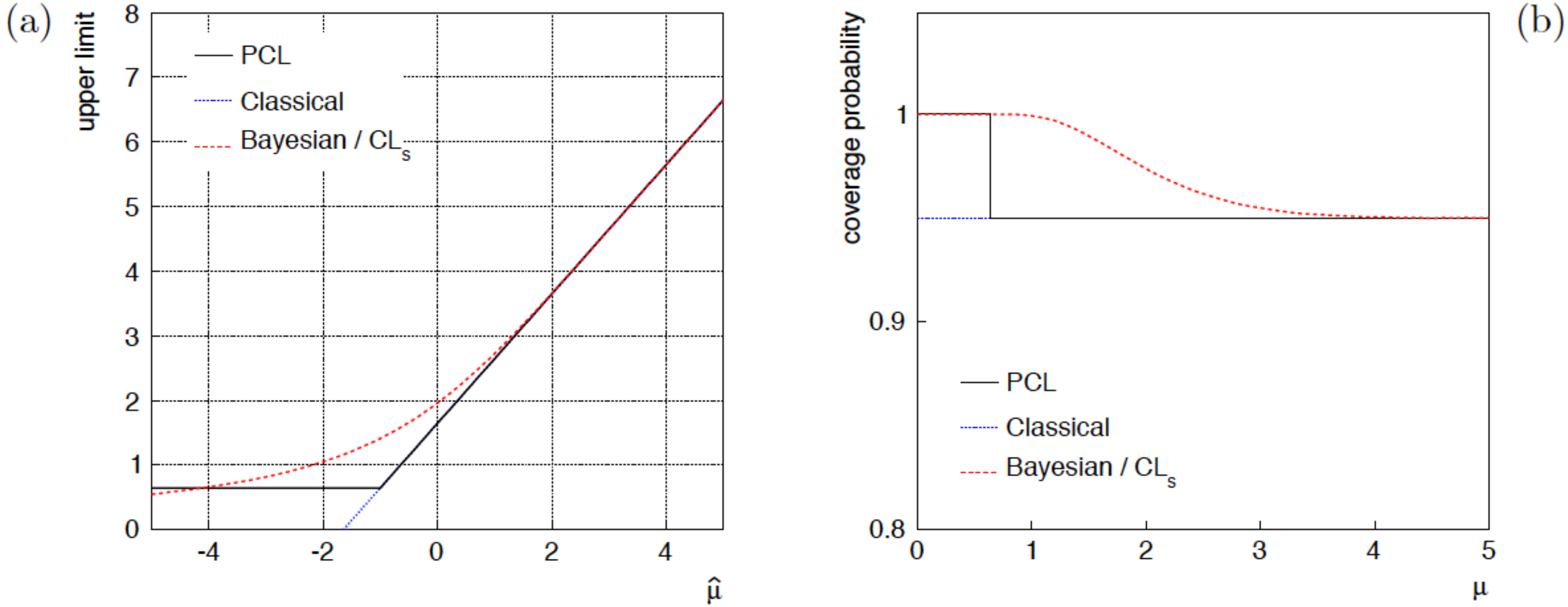
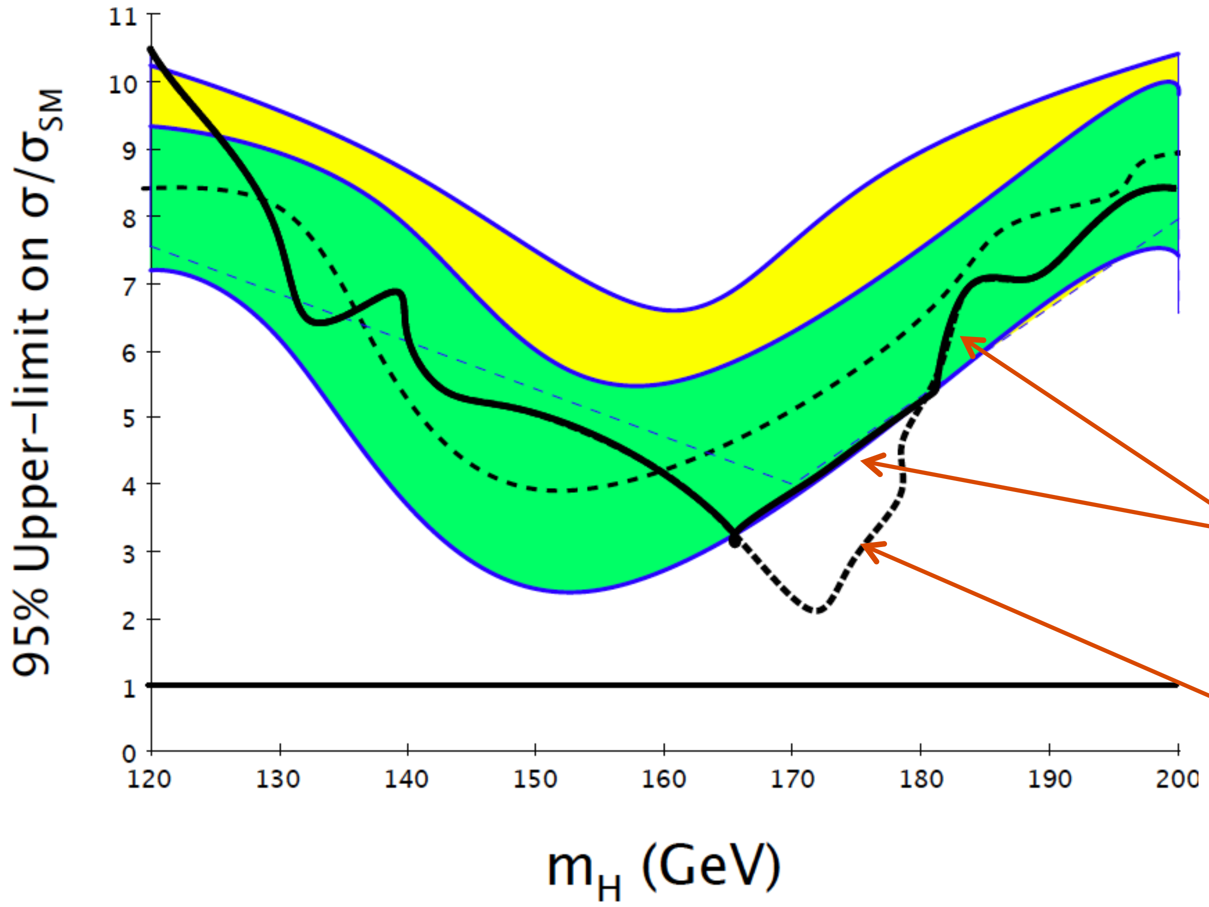


Figure 3: (a) Upper limits from the PCL (solid), CL_s and Bayesian (dashed), and classical (dotted) procedures as a function of $\hat{\mu}$, which is assumed to follow a Gaussian distribution with unit standard deviation. (b) The corresponding coverage probabilities as a function of μ.

PCL as a function of, e.g., m_H



PCL

Here power below threshold; do not exclude.

Some reasons to consider increasing M_{min}

M_{min} is supposed to be "substantially" greater than α (5%).

So M_{min} = 16% is fine for $1 - \alpha$ = 95%, but if we ever want $1 - \alpha$ = 90%, then16% is not "large" compared to 10%; $\mu_{\text{min}} = 0.28\sigma$ starts to look small relative to the intrinsic resolution of the measurement. Not an issue if we stick to 95% CL.

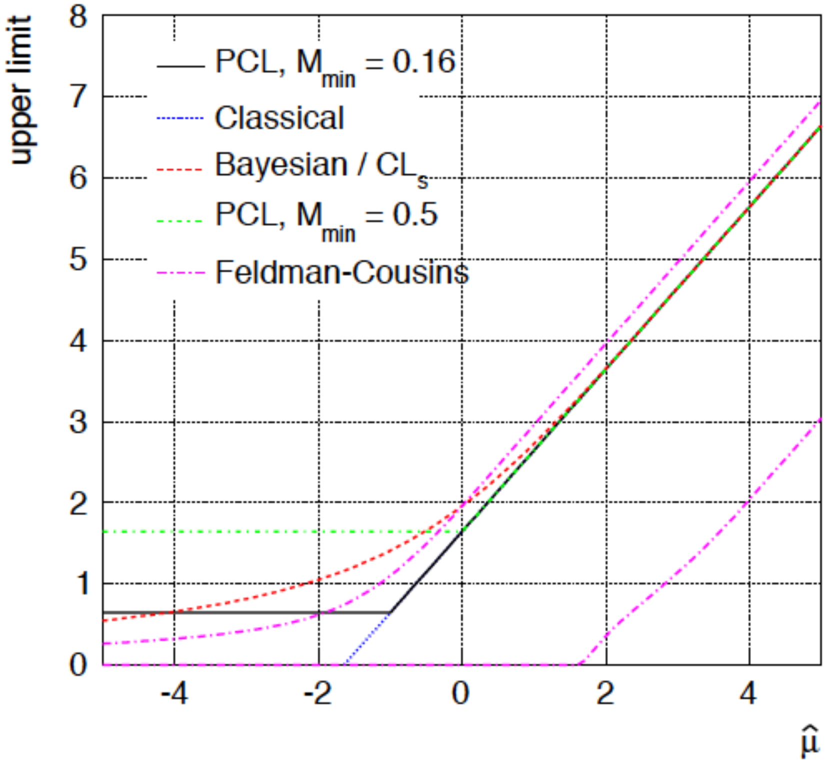
PCL with M_{min} = 16% is often substantially lower than CLs. This is because of the conservatism of CLs (see coverage).

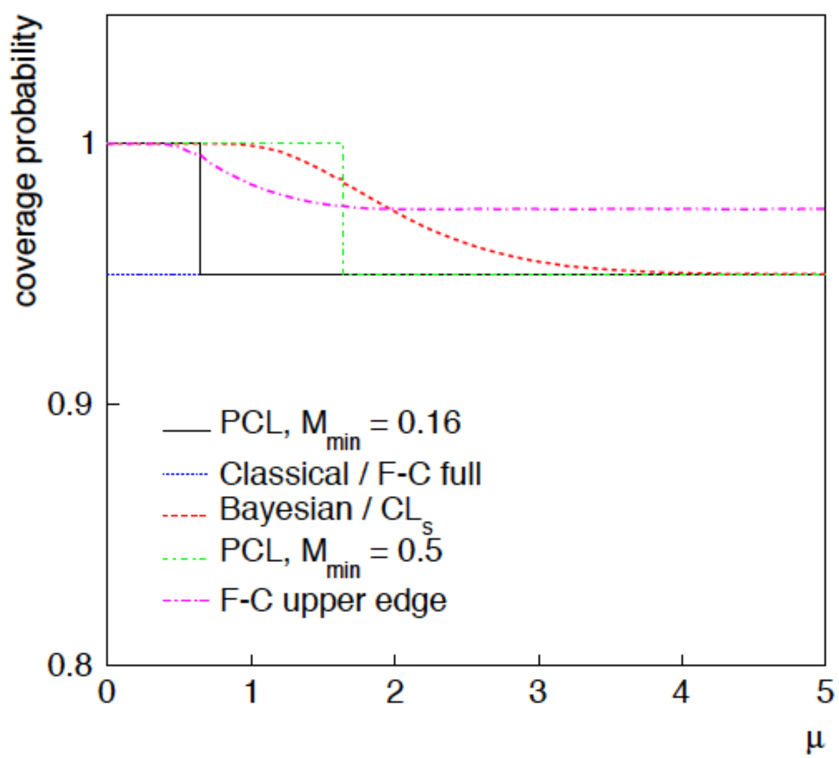
But goal is not to get a lower limit per se, rather

(a) to use a test with higher power in those regions where one feels there is enough sensitivity to justify exclusion and

(b) to allow for easy communication of coverage (95% for $\mu \geq \mu_{\text{min}}$; 100% otherwise).

PCL with M_{min} = 0.16, 0.50 (and other limits)





With M_{min} = 50%, power constraint is applied half the time.

This is somewhat contrary to the original spirit of preventing a "lucky" fluctuation from leading to a limit that is small compared to the intrinsic resolution of the measurement.

But PCL still lower than CLs most of the time (e.g., $x > -0.4$).

Aggressive conservatism

It could be that owing to practical constraints, certain systematic uncertainties are over-estimated in an analysis; this could be justified by wanting to be conservative.

The consequence of this will be that the +/-1 sigma bands of the unconstrained limit are broader than they otherwise would be.

If the unconstrained limit fluctuates low, it could be that the PCL limit, constrained at the -1sigma band, is lower than it would be had the systematics been estimated correctly.

conservative = aggressive

If the power constraint M_{min} is at 50%, then by inflating the systematics the median of the unconstrained limit is expected to move less (or not at all).

A few further considerations

Obtaining PCL requires the distribution of unconstrained limits, from which one finds the M_{min} (16%, 50%) percentile.

In some analyses this can entail calculational issues that are expected to be less problematic for M_{min} = 50% than for 16%.

Analysts produce anyway the median limit, even in absence of the error bands, so with M_{min} = 50% the burden on the analyst is reduced somewhat (but one would still want the error bands).

Changing M_{min} now may be too short notice for EPS, or, one may argue that if we are going to make the change, now is the time.

We propose moving M_{min} to 50% and to report both PCL and CLs.

We propose using PCL as the primary result.