Preferred measure of central tendency

NOAA FISHERIES
Alaska Fisheries Science Center

Grant Thompson

November 13, 2017

Background (1 of 3)

- SSC has long been interested in model averaging for EBS Pcod
- 12/08: "Consider the strengths and weaknesses of model averaging as an alternative to model selection and provide a rationale for or against use of this method in future assessments."
- 10/16: "The observed discrepancies among different models in these assessments are a good – if perhaps extreme – example of the model uncertainty that pervades most assessments. This uncertainty is largely ignored once a model is approved for specifications. We encourage the authors and Plan Teams to consider approaches such as multi-model inference to account for at least some of the structural uncertainty."
- ...and many other comments that did not make it into the minutes

Background (2 of 3)

- In the 2016 EBS Pcod assessment, the author suggested: “As an appropriate method for using a full model averaging approach in the context of the current management framework has yet to be determined, a possible short-term compromise would be to choose the single model that gives a 2017 maximum permissible ABC closest to the average across all models.”
- 12/16: “The SSC noted that choosing a model that is somewhere ‘in the middle’ of the set is not a good approach to model averaging as it ignores within-model uncertainty (by using a naïve average of the point estimates)”
- But wait--the mean of the averaged distribution is just the average of the means of the individual distributions (stay tuned!)

Background (3 of 3)

- 12/16: “The SSC [recommends] further considering model averaging based on the outcome of the SSC workshop during the February 2017 meeting.”
- 2/17: “The SSC would like to see a ‘test case’ of how ensemble modeling works for one of our groundfish stocks.”
- Debate among SSC members during the 2/17 workshop: Is the central tendency of the averaged distribution equal to the average of the central tendencies of the individual distributions?
- In an email discussion with an SSC member during the September BSAI Team meeting, the misunderstanding became clear: Some people were thinking of the *mean* as the central tendency, while others were thinking of the *median* as the central tendency

SSC request (directed to EBS Pcod author)

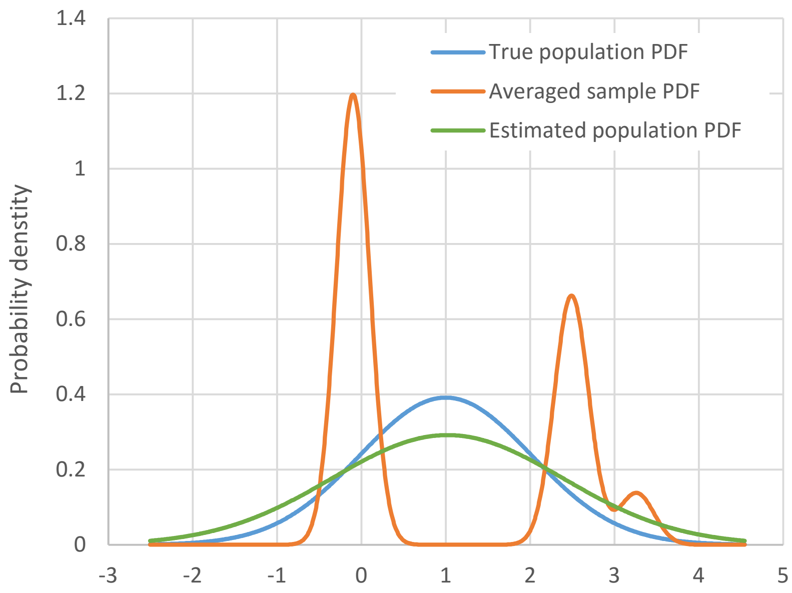
- 10/17: “Clarify, with the Joint Plan Teams, the preferred measure of central tendency (e.g., median or mean) for assessments reporting probabilistic results either via Bayesian posteriors or model-averaged distributions.”
- Note: In the interest of efficiency, “average” as used in this presentation can mean either “weighted average” or “unweighted average” (same as “equally weighted average”)
  - This is not a presentation on model weighting

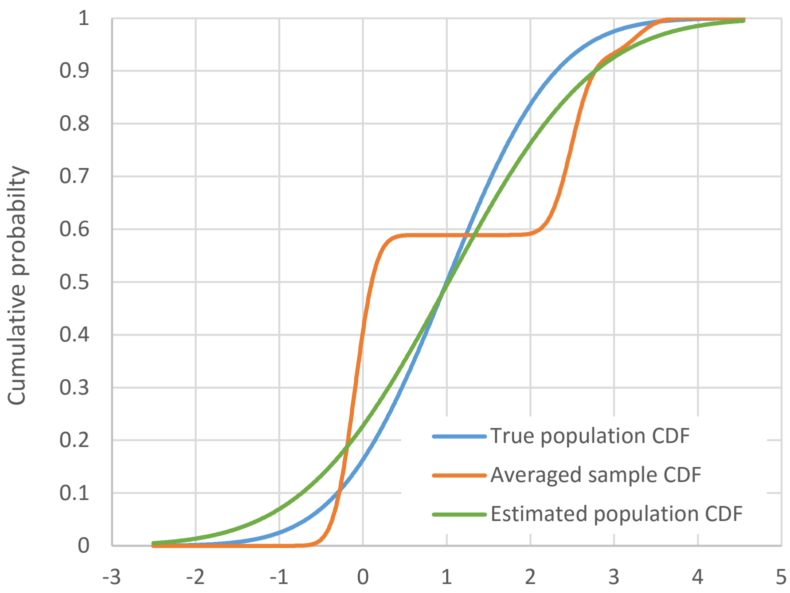
Approaches to alternative models

1. From the population of all possible models, examine only 1
2. From the population of all possible models, examine a sample of size > 1
   A. Use only 1 model in the sample
   B. Use all models in the sample by assigning non-zero weights to each
      a. Assign a weight of 0 to all models *not* in the sample
      b. Use the averaged sample distribution to estimate the population distribution
         - Method is just exploratory at this point
         - EBS Pcod assessment provided it as an option

Hypothetical example

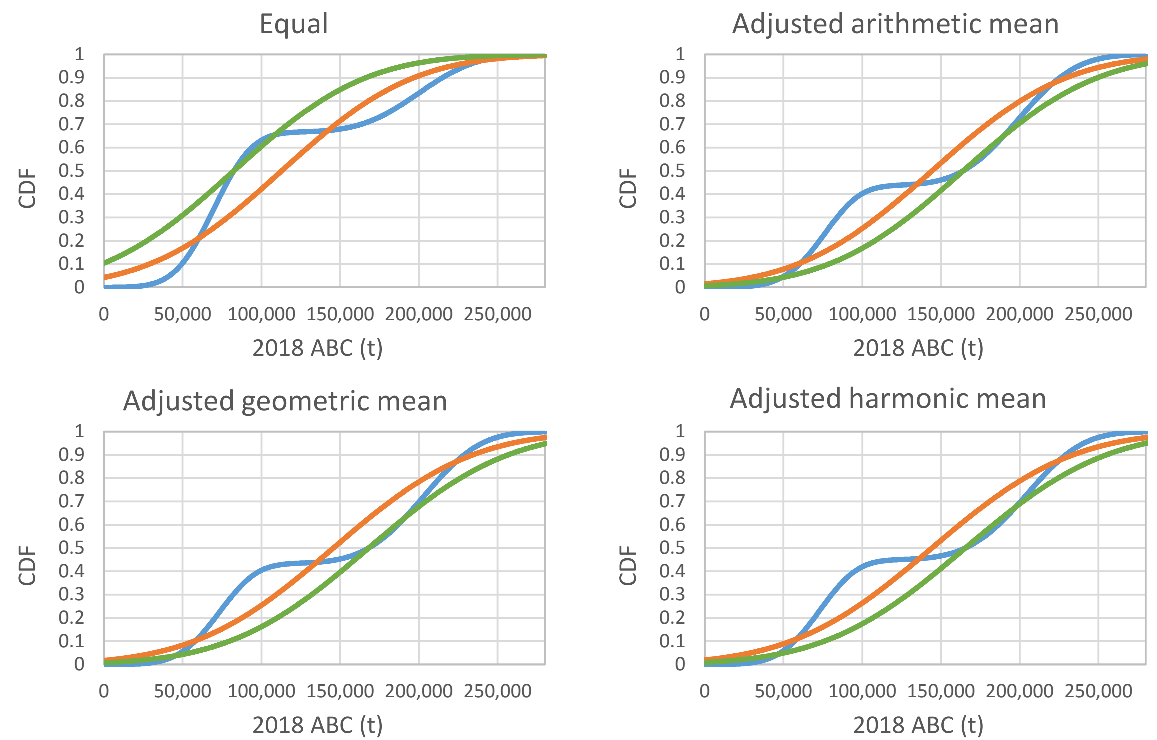
- Distributions of μ and σ (for normal distribution) given
- “True” population distribution integrated over μ and σ
- Three models with μ and σ drawn at random, weighted “correctly”
- “Estimated” population distribution based on moments from sample





Actual example: EBS Pcod assessment

- Blue = sample distribution, orange = population distribution fit to mean, green = population distribution fit to median
- No way of knowing the true population distribution, of course



Fitting the true population distribution

- Which fits the true population distribution better: the averaged sample distribution or the estimated population distribution?
- 10,000 simulations conducted each for n=5 and n=10 (models)
- Unimodal "true" population distribution
- Goodness of fit measured by Kullback-Leibler divergence
- Frequency with which the estimated population distribution gave the better fit varied with both sample size and number of modes

	No. models in sample = 5		No. models in sample = 10	
Modes	No. sims.	Est. pop. wins (%)	No. sims.	Est. pop. wins (%)
1	6646	79.8%	7819	99.9%
2	1622	92.2%	1054	99.8%
3	716	96.6%	364	100.0%
4	292	99.7%	93	100.0%
5	60	95.0%	10	90.0%
All:	9336	84.0%	9340	99.9%

Approaches to uncertainty

1. Frequentist approach
   - Example: P* approach to setting ABC
     - Given the distribution of the true-but-unknown OFL, set ABC such that the CDF evaluated at ABC equals P*
   - Percentiles (e.g., the median if P*=1) are key
2. Bayesian ("decision-theoretic") approach
   - Example: constant relative risk aversion (RRA)
     - Given the distribution of long-term yield conditional on FABC, set FABC so as to maximize the mean of order 1-RRA
   - Expected values (e.g., the mean if RRA=0) are key

Properties of sample mean and median

- The sample mean is an unbiased estimator of the population mean, and the sample median is an unbiased estimator of the population median
- In general, the sample median has a larger variance than the sample mean
  - E.g., if the population distribution is normal, the variance of the sample median will be greater than the variance of the sample mean (asymptotically) by a factor of $\pi/2$
- If the population distribution is symmetric, the population mean and median are equal, in which case the sample mean is a better estimator (than the sample median) of either the population mean or median

Suggested policy: “It depends”

- Frequentist approaches naturally lend themselves to use of percentiles, such as the median
  - If the “final” distribution is just the averaged sample distribution, then the median is a more useful choice than the mean
  - If the “final” distribution is the population distribution as inferred from the statistics of the averaged sample distribution, the sample mean will provide a better estimate of the population distribution and its various percentiles, such as the median
- Bayesian approaches naturally lend themselves to use of moments, such as the mean
  - Regardless of sample distribution vs. population distribution
- Of course, if it is just a matter of *reporting*, easy enough to do both