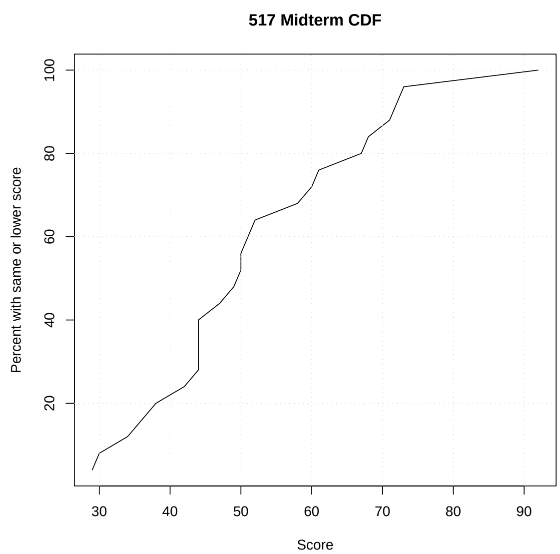
## FREQUENTLY OCCURRING ERRORS ON THE MIDTERM

The midterm exam grades had a shorter lower tail than usual — few people far behind the median score. The lowest maximum score over all questions was 14 out of 23, so everyone had at least one area of demonstrated competence. No letter grades are attached to the scores, since these will simply be added in numerically to scores on the final and the problems sets. The A/A- range would probably be around 60 or 65 and up, the B+/B- range around 35 or 40 up, and scores below that in the C range. None of the midterm scores are below the C range. The cdf of exam scores is shown in the figure.




Here are some mistakes that occurred on multiple exams and that represent important misunderstandings.

**Problem 1:**

- There were assertions that the sample mean couldn't be a Bayesian decision rule because, since the parameter $p$ can take on only discrete values, the Bayesian rule would also have to take on only discrete values. But with squared-error loss functions (as in this problem), the Bayesian decision rule is the posterior mean, which will generally not be discrete.
- The "deadweight loss" C is not the decision-theoretic "loss" in this problem. Some people calculated a deadweight loss table in part (a), but then went on to recognize that the decision problem maximizes B-C

©2004 by Christopher A. Sims. This document may be reproduced for educational and research purposes, so long as the copies contain this notice and are retained for personal use or distributed free.

(or minimizes "loss" C-B). But a few continued as if deadweight loss from taxes were the only element of the decision-theoretic loss function, which makes the problem trivial.

- The definition of admissibility makes no mention of probabilities. An estimator that fails to be optimal no matter what probability distribution is assigned to states is not necessarily inadmissible, and we had examples of that in an exercise. Similarly an estimator that is optimal for some probability distribution is not necessarily admissible. These are instead the definitions of estimators that are not, and are, Bayesian decision rules. The Bayesian rules are often nearly or exactly identical to the class of admissible rules, but there are borderline exceptions, as in this problem, where one decision rule was inadmissible, but nonetheless equivalent in expected loss to the optimum rule under a certain distribution.

**Problem 2:**

- In finding the likelihood or posterior kernel, it is safe to drop factors that depend on data but not on the parameters, but not to drop factors that depend on parameters but not data. In this problem, the $\Gamma(p)$ "normalizing constant" in the pdf cannot be dropped.
- Several exam papers, and ones I have seen on final and general exams, seem to have trouble with converting a single-observation pdf to a sample pdf. The posterior is written down as "likelihood times prior", with a single observation's pdf written out as if it were the likelihood. A variant in this problem was to write the likelihood with $x^T$ replacing the correct $\prod_1^T x_t$ in the likelihood.

**Problem 3:**

- The sum of two independent $U(0, 1)$ variates is not distributed as $U(0, 2a)$.
- The acceptance region for a test can often be written as $\beta \pm \delta$, with $\hat{\beta}$'s outside that region leading to rejection. This is an interval of observed $\beta$ values, functions of the data. A confidence interval is an interval of unobserved true parameter values $\beta$, and will often have the form $\hat{\beta} \pm \delta$. In SNLM cases and some similar models, the $\delta$ can be the same, so not being careful about $\hat{\beta}$'s and $\beta$'s may not be a problem. But in this question's setup, because the length of the interval varied with the parameter, the CI and acceptance region were very different.
- The likelihood is $p(X \mid \beta)$ *as a function of* $\beta$. Quite a few plots of $p(X \mid \beta)$ as a function of $X$ appeared as responses to the request for a sketch of the likelihood function.

**Problem 4:**

- This one is probably just a reflection of stress, but the several people who wrote down formulas with $\sqrt{s^2(X'X)^{-1}}$ entering as if it were a well-defined scalar should review their linear algebra if they don't understand why that formula makes no sense.
- Part (a) of this question asked for a confidence interval. No prior or posterior pdf's are involved, because confidence sets are a non-Bayesian notion. Several exam papers began this question by assuming a prior and describing the posterior.
- Though most answers to (a) used the fact that a linear combination of multivariate $t$'s is univariate $t$, a few constructed an $F(1,7)$ pivot. This pivot is distributed as the square of a $t_7$. But if one constructs a two-sided interval for the $F$ pivot, this corresponds to a disconnected pair of intervals for $\beta_1 + \beta_2$, since it rejects for both small and large values of the $t$ statistic.
- In part (b), to find the required interval by simulation one has to draw $\beta_1$ and $\beta_2$ jointly, because they are not independent. Several exam papers suggested that because we know each is a $t_7$ variate, we could draw them one at a time from the appropriate $t$'s.