APPROACHES TO INFERENCE EXERCISE

In class we considered these two samples:

A

T=1	5	1.1	1.1	1.1	1.1
T=0	.9	.9	.9	.9	.9

B

T=1	5	.9	.9	.9	.9
T=0	1.1	1.1	1.1	1.1	1.1

The difference in means between treated and untreated is .98 for sample A and .61 for sample B. We observed that a simple randomization test, of the null that the observed data in the table has been generated by randomly allocating the 10 observed numbers in the table across the 10 positions in the table, rejects the null at marginal significance level .004 for sample A but accepts it at marginal significance level .5 for sample B. A model for the two samples that treated them both as drawn i.i.d. from some distribution with continuous density would be unlikely to generate such sharply contrasting results for samples, like these, close to each other in Euclidean distance.

Calculate the posterior probability that the model from which the $T = 1$ data are drawn has a higher mean than the $T = 0$ data for sample A and sample B, under these assumptions

(1) $T = 1$ and $T = 0$ data are both drawn from exponential distributions (i.e. with densities ae^{-ay}), with any differences in means arising from differences between the values of a between the $T = 0$ and $T = 1$ data. Assume independence, conditional on T and the parameters, among all draws of the data. Assume a flat prior. Note that under these assumptions the posterior pdf for a in each sample is in the shape of a Gamma distribution, and therefore can be drawn from directly by commands in most matrix programming languages (e.g. `rgamma(M, n, rate)` in R). Evaluating the integral over the parameter space required here analytically is possible in principle, though probably not worth it, since drawing, say, 1000 random numbers from the posteriors for the a values for $T = 0$ and $T = 1$ is a single command.

(2) Assume asymptotic normality applies here despite being applied to samples of size 5. That is, calculate the sample mean and standard deviation for $T = 1$ and for $T = 0$. Let these four objects be m_1, s_1, m_0, s_0. s_0 is zero in these data, which leads to treating m_0 as if it were a non-random constant, but m_1 asymptotically normal with variance $s_1^2/5$ This implies μ_1, the true value of the mean for $T = 1$ has a normal posterior distribution. You can then calculate analytically the posterior probability that $\mu_1 > \mu_0 = m_0$.

Date: February 16, 2020.

(3) Extra credit: Assume that the $T = 0$ and $T = 1$ samples are each from a normal distribution with unknown mean and variance, and use a conjugate prior on the means and variances. For example, assume an inverse-gamma prior with shape parameter 1, scale 1 (IG(1,1)) for σ^2 in each sample and a $N(0, 5\sigma^2)$ for $\mu \mid \sigma^2$. This is a quite flat prior and is conjugate, so it is possible to draw directly from the implied posterior. This is a lot of work to do for 10 observations, so it's not required. Nonetheless it seems that allowing for uncertainty about σ^2 could make inference more conservative, and it would be interesting to find out.