Econ 8013 Inference using summary statistics and comparison of means Jeff Mills, March 2017

Given a table of summary statistics from a trial, such as Table 1, it is possible to perform Bayesian inference to compare the distribution of the mean outcome. ** **See p.107 Hahn (2014) for two-sample comparison example.**

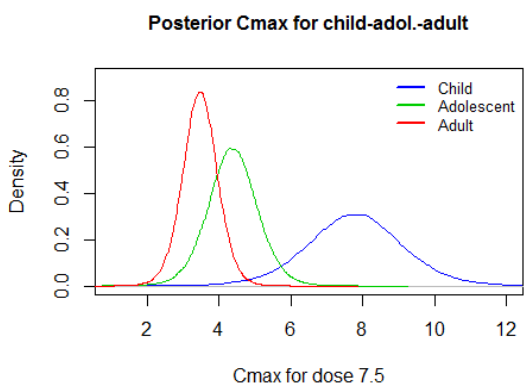
Table 1

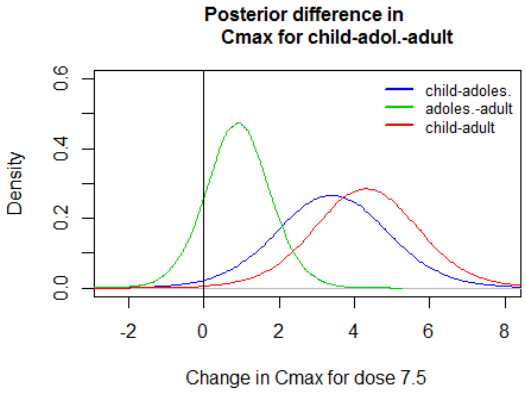
Dose	Cmax Child N = 12	Cmax Child SD	Cmax Adoles. N = 12	Cmax Adoles. SD	Cmax Adult N = 14	Cmax Adult SD
7.5	7.78	4.38	4.38	2.25	3.47	1.74
15	14.19	5.91	8.68	4.02	7.19	3.01
30	22.83	5.62	17.12	6.97	13.98	6.08

Following the procedure described below, we obtain the following results [see **Cmax_1PPresults.R** - uses **postdiff3.R** – function to generate MC sample from three Student-t posteriors]

Steady State Cmax 1-PP comparison across child-adoles.-adult

Dose = 7.5





p-values and odds vs. 0

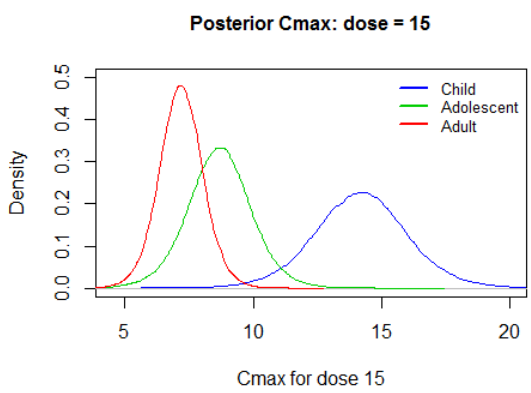
```
[1,] "       "    " Child-Adoles." " Adoles.-Adult" "Child-Adult"
[2,] "p-value"    "0.052"          "0.324"          "0.012"
[3,] "odds vs. 0" "8.56"           "1.73"           "32.66"
```

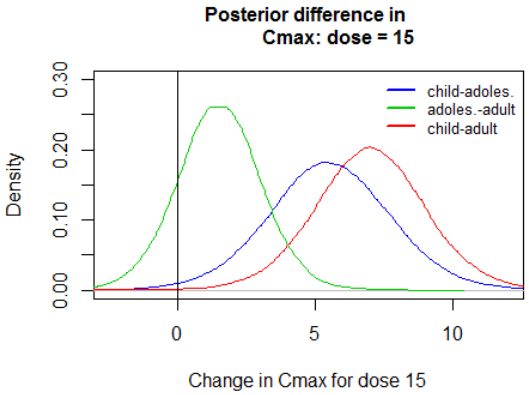
Means, SDs, and credible intervals for difference

```
  mean(ds1s2)= 3.407746, sd(ds1s2)= 1.566315
> quantile(ds1s2,probs=c(0.005,0.025,0.5,0.975,0.995))
      0.5%       2.5%        50%      97.5%      99.5%
-0.8772345  0.3222240  3.4041057  6.5134228  7.7250151

  mean(ds2s3)= 0.905402, sd(ds2s3) = 0.8790466
> quantile(ds2s3,probs=c(0.005,0.025,0.5,0.975,0.995))
      0.5%       2.5%        50%      97.5%      99.5%
-1.4751838 -0.8335557  0.9058538  2.6457158  3.2664697
```

```
> cbind(mean(ds1s3), sd(ds1s3))
 [1,] 4.313148 1.477931
> quantile(ds1s3,probs=c(0.005,0.025,0.5,0.975,0.995))
     0.5%      2.5%       50%     97.5%     99.5%
0.2416432 1.3810908 4.3119458 7.2465036 8.4072141
```

Dose = 15





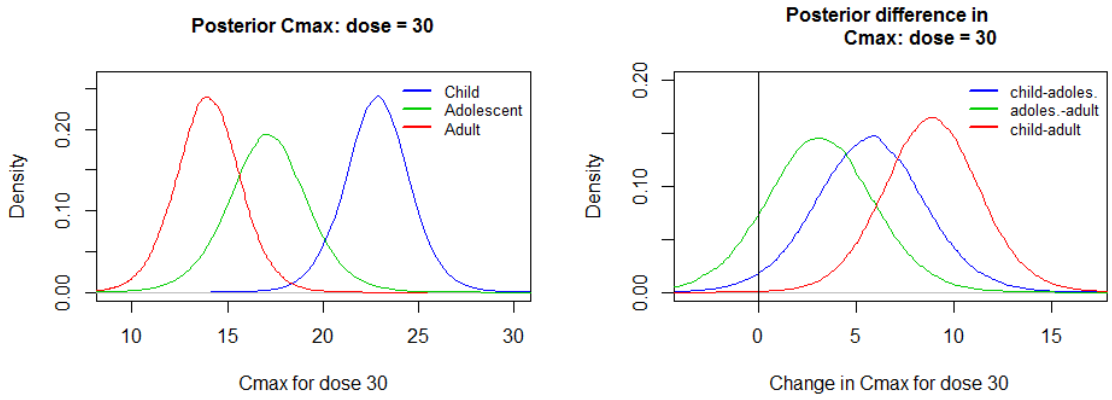
p-values and odds vs. 0

```
[1,] "       "    " Child-Adoles." " Adoles.-Adult" "Child-Adult"
[2,] "p-value"    "0.034"          "0.358"          "0.006"
[3,] "odds vs. 0" "12.83"          "1.61"           "75.8"
```

Means, SDs, and credible intervals for difference

```
> cbind(mean(ds1s2), sd(ds1s2))
 [1,] 5.508496 2.281496
> quantile(ds1s2,probs=c(0.005,0.025,0.5,0.975,0.995))
      0.5%       2.5%        50%      97.5%      99.5%
-0.6352811  0.9888739  5.4952653 10.0167327 11.7099625

> cbind(mean(ds2s3), sd(ds2s3))
 [1,] 1.487398 1.554897
> quantile(ds2s3,probs=c(0.005,0.025,0.5,0.975,0.995))
     0.5%      2.5%       50%     97.5%     99.5%
-2.750811 -1.595043  1.488953  4.538548  5.718716

> cbind(mean(ds1s3), sd(ds1s3))
 [1,] 6.995894 2.076199
> quantile(ds1s3,probs=c(0.005,0.025,0.5,0.975,0.995))
     0.5%      2.5%       50%     97.5%     99.5%
 1.305385  2.885799  6.994985 11.112773 12.687794
```

Dose = 30



p-values and odds vs. 0

```
[1,] "          "    " Child-Adoles." " Adoles.-Adult" "Child-Adult"
[2,] "p-value"    "0.07"           "0.29"           "0.004"
[3,] "odds vs. 0" "6.47"           "1.88"           "100.15"
```

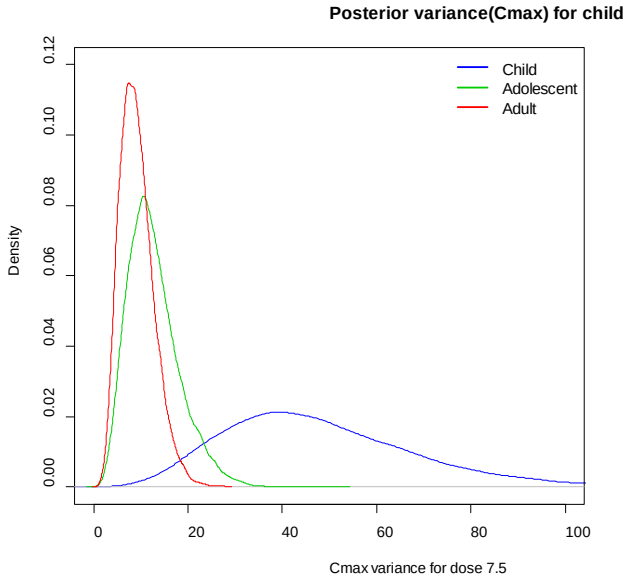
Means, SDs, and credible intervals for difference

```
> cbind(mean(ds1s2), sd(ds1s2))
 [1,] 5.717711 2.85352
> quantile(ds1s2,probs=c(0.005,0.025,0.5,0.975,0.995))
       0.5%        2.5%         50%       97.5%       99.5%
-1.97733493  0.07329689  5.72373566 11.38740364 13.45472066

> cbind(mean(ds2s3), sd(ds2s3))
 [1,] 3.137964 2.837423
> quantile(ds2s3,probs=c(0.005,0.025,0.5,0.975,0.995))
     0.5%      2.5%       50%     97.5%     99.5%
-4.516691 -2.493728  3.137395  8.745445 10.830198

> cbind(mean(ds1s3), sd(ds1s3))
 [1,] 8.855675 2.51722
> quantile(ds1s3,probs=c(0.005,0.025,0.5,0.975,0.995))
     0.5%      2.5%       50%     97.5%     99.5%
 2.092619  3.903527  8.852022 13.861130 15.698965
```

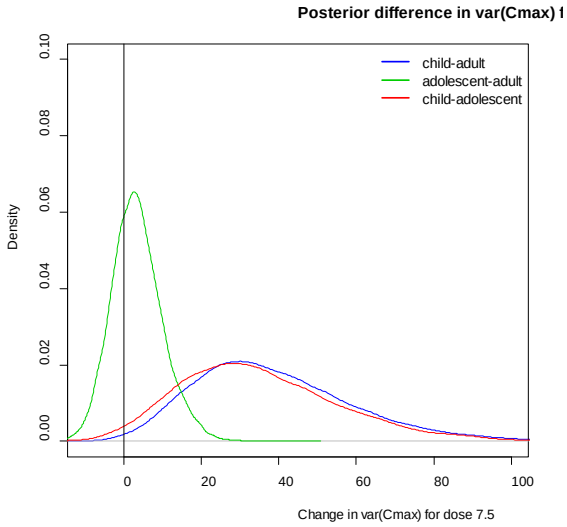
Code:
Cmax_1PPresults.R – analysis of Cmax 1PP data
Half-life_1PPresults.R – analysis of half-life 1PP data
postdiff3.R – function to generate MC sample from three Student-t posteriors

Variance

We can also looks at the variance; something that is difficult to do using frequentist methods.



Large variance for child dosage makes sense, since relatively small changes in age of the child can make a big difference.




```
 quantile(d7515,probs=c(0.005,0.025,0.5,0.975,0.995))
      0.5%       2.5%        50%      97.5%      99.5%
 -5.348499   1.529884  32.691370  83.445918 102.453525

 quantile(d1530,probs=c(0.005,0.025,0.5,0.975,0.995))
      0.5%       2.5%        50%      97.5%      99.5%
-12.177943  -8.287523   3.079146  17.752620  23.116933

 quantile(d7530,probs=c(0.005,0.025,0.5,0.975,0.995))
       0.5%         2.5%          50%        97.5%        99.5%
 -0.04210567   6.54397400  35.88186376  86.34139047 104.00262019
```

Summary

Suppose $x \sim N(\mu, \tau)$, where the precision. τ, is the inverse of the variance, σ^2, i.e. $\tau = 1/\sigma^2$. Given a sample, x, and adopting an uninformative prior, we have a Student-t marginal posterior for μ,

$$\mu \sim t(\bar{x}, s^2, \nu),$$

and a Gamma marginal posterior for τ,

$$\tau \sim Gamma\left(\frac{\nu}{2}, \frac{\nu s^2}{2}\right),$$

with the summary statistics $\bar{x}$, s^2 and ν sufficient information to completely determine these distributions. Note that the marginal posterior for σ^2 is therefore Inverted Gamma with the same parameters as for τ, i.e. $\sigma^2 \sim IG(\nu/2, \nu s^2/2)$.

Theory

Given only summary statistics, it is necessary to make a distributional assumption. An assumption of normally distributed outcomes appears reasonable from experience with data from other similar experiments. This leads to the likelihood function,

$$p(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}\right),$$

where $x = (x_1, x_2, \ldots, x_n)$, x_i = observed C_{max} for individual i. It is mathematically convenient to reparameterize in terms of the precision $\tau = 1/\sigma^2$ (the inverse of the variance), so this becomes,

$$p(x|\mu, \tau) \propto \tau^{\frac{n}{2}} \exp\left(-(\tau/2)\sum_{i=1}^{n}(x_i - \mu)^2\right),$$

where $\propto$ indicates proportionality, i.e. 'equal to up to a constant of proportionality', so that $x \propto y$ is equivalent to $x = cy$, c a constant.

Under the normality assumption, the sample mean, sample standard deviation, and number of observations are sufficient statistics to fully determine the likelihood, so Bayesian inference is then possible.

The conventional vague, uniform prior for μ and $\log(\tau)$, namely

$$p(\mu, \tau) \propto \frac{1}{\tau}, \quad -\infty < \mu < \infty, \quad \tau > 0,$$

is a standard choice to represent uninformative prior beliefs concerning μ and τ.

This leads to the joint posterior density,

$$p(\mu, \tau|x) \propto \tau^{\frac{n}{2}-1} \exp\left(-(\tau/2)\sum_{i=1}^{n}(x_i - \mu)^2\right).$$

Integrating w.r.t τ leads to the marginal posterior density for μ,

$$p(\mu|x) \propto \left(1 + \frac{n(\mu - \bar{x})^2}{\nu s^2}\right)^{-\frac{n}{2}},$$

which is a Student-t density with $\nu = n - 1$, $\bar{x} = \sum x_i/n$, and $s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/\nu$.

Marginalizing w.r.t to μ gives the marginal density for τ (or σ^2 if we substitute for $\tau = 1/\sigma^2$),

$$p(\tau|x) \propto \tau^{\frac{\nu}{2}-1} \exp\left(-\frac{\tau \nu s^2}{2}\right),$$

which is a Gamma density. [see, for example, Lancaster (2004), p. 120-125]

The marginal density for μ is generally the posterior of interest. Given two sample means we wish to compare, we can generate a large Monte Carlo (MC) pseudo-sample by obtaining random draws from this Student-t distribution, then compute the differences to obtain the posterior density of the difference in means for two samples.

With an MC sample from the posterior, the accuracy can be made arbitrarily high by increase the pseudo-sample size. The law of large numbers then assures that the expected value of any function of the MC sample converges its true value, i.e. for pseudo-sample of R draws for z, as $R \to \infty$,

$$E\big(f(z)\big) \to \frac{1}{R}\sum_{r=1}^{R} f(z^{(R)}),$$

where $z^{(r)}$ is the rth pseudo-sample draw.

The algorithm for comparison of means is then as follows.

1. Given sample means, $\bar{x}_1$ and $\bar{x}_2$, sample standard deviations s_1 and s_2, and samples sizes n_1 and n_2, randomly draw R values, μ_j^r, $r = 1,2, \ldots, R$, from each of the marginal posterior distributions $p(\mu_1|x_1)$ and $p(\mu_1|x_1)$, which are Student-t distributions as given above.

2. Compute R differences in means, $\Delta\mu^{(r)} = \mu_1^{(r)} - \mu_2^{(r)}$.

3. Quantiles, means, etc. can then be computed from the sample of differences in means.

Variances can be compared in the same way by drawing from the Gamma distribution.

Drawing from Gamma in R: The Gamma distribution with parameters shape = *a* and scale = *b* has density

$$p(z|a, b) \propto z^{a-1} \exp(-bz)$$

We have

$$p(\tau|x) \propto \tau^{\frac{\nu}{2}-1} \exp\left(-\frac{\tau\nu s^2}{2}\right),$$

so, $s = \nu s^2/2$ and $a = \nu/2$.

To generate R draws from this distribution in R we use: rgamma(R,shape = a, scale = b).

Sequential analysis via Bayesian updating

We can also examine different data samples sequentially, updating our posterior inference given a new sample. As above, suppose we have a sample $x \sim N(\mu, \tau)$. Adopting an uninformative prior, we have a Student-t marginal posterior for μ, $\mu \sim t(\bar{x}, s^2, \nu)$, and a Gamma marginal posterior for τ, $\tau \sim Gamma(\nu/2, \nu s^2/2)$ [or $\sigma^2 \sim IG(\nu/2, \nu s^2/2)$], with the summary statistics $\bar{x}$, s^2 and ν.

Suppose now that a second sample of m observations is obtained from the same distribution, $y \sim N(\mu, \tau)$. We can use the posterior distributions obtained above with the first sample as the new prior distribution. Combining with the likelihood for y, it can be shown [see below] that the marginal posteriors based on the combined sample $z = (x, y)$ are a Student-t marginal posterior for μ,

$$\mu \sim t(\bar{z}, s_z^2, \nu_z),$$

and a Gamma marginal posterior for τ

$$\tau \sim Gamma(\nu_z/2, \nu_z s_z^2/2),$$

with the summary statistics,

$$\bar{z} = \frac{n\bar{x} + m\bar{y}}{n+m},$$

$$s_z^2 = \frac{ns^2 + ms_y^2}{n+m}, \qquad (*)$$

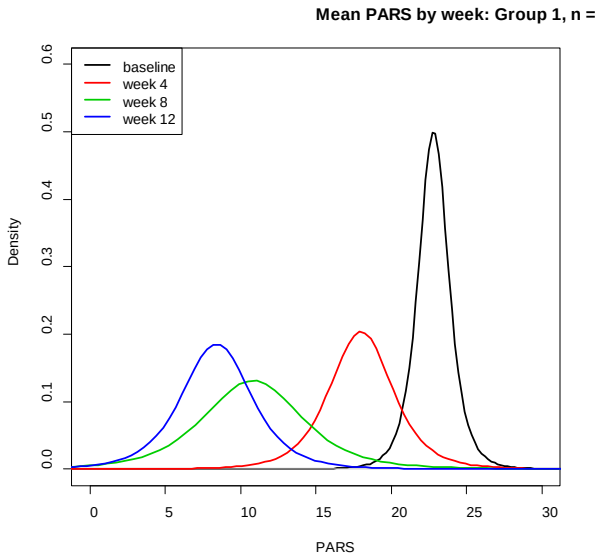
$$\nu_z = \nu + m,$$

with $s_y^2 = \sum_{i=1}^{m} (y_i - \bar{y})^2 / m$.

The above summary statistics provide sufficient information to completely determine these distributions, so these updating equations can be used as more observations become available and the functional forms of the posteriors remain the same.

Given a third sample, $w \sim N(\mu, \tau)$, the above posteriors can be used as the new prior so that the same updating equations (*) can be used. This process can be iterated for any number of samples, providing sequential updating of the posterior as new sample information becomes available.

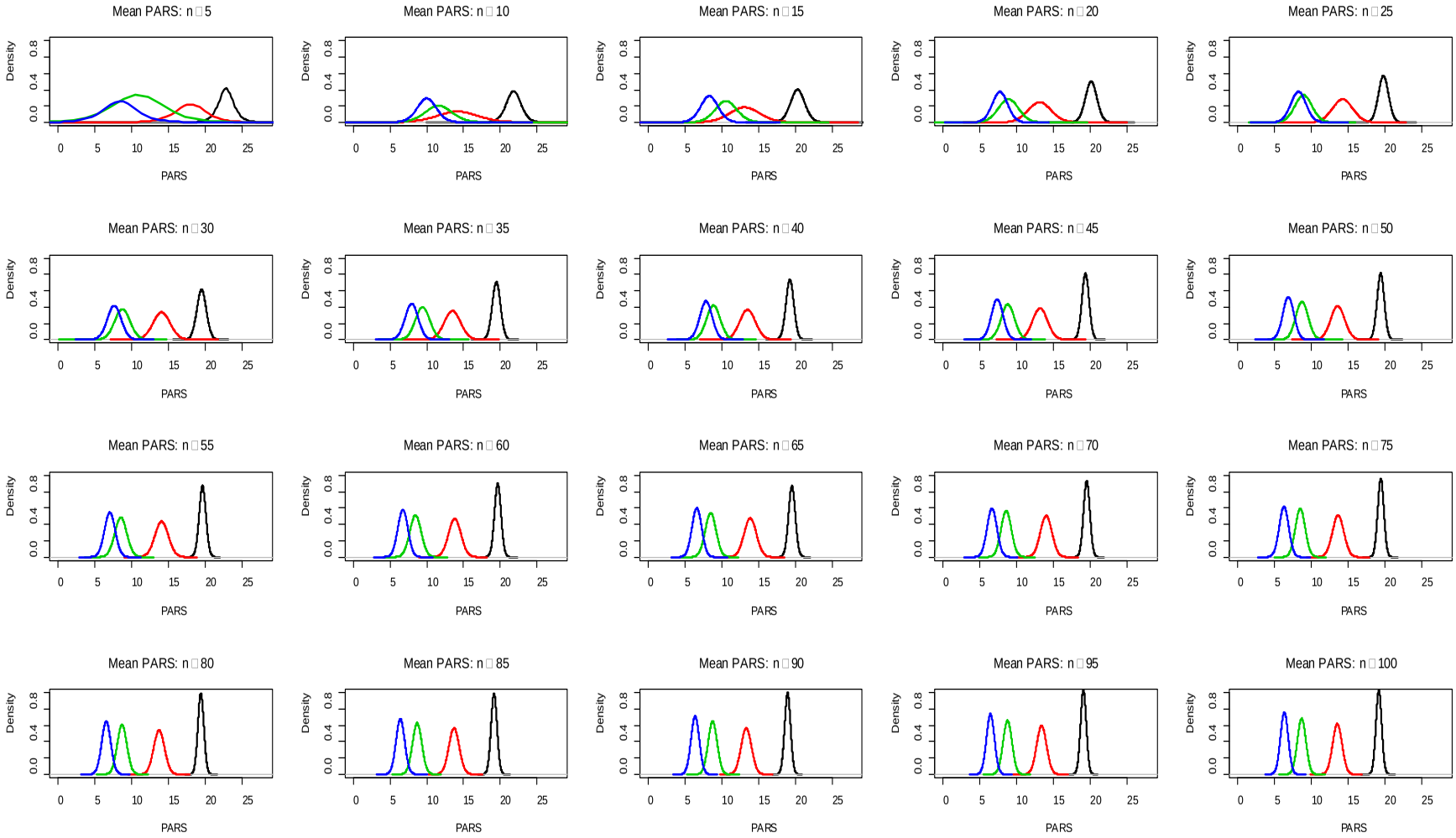
For comparison of treatment vs. nontreatment (or placebo), we obtain the posteriors for the means of each group, then we can simulate to obtain a pseudo-sample from each exact posterior distribution. These MC samples can then be combined to numerically obtain the posterior distribution of the difference in means for inference and hypothesis testing.



Sequential analysis of PARS treatment group 1: summary statistics for data

```
   means  mw0     mw4    mw8     mw12
          22.8     18    10.8    8.4
SDs       sdw0    sdw4    sdw8    sdw12
         2.280   4.743   7.429   5.272
```

Sequential posteriors adding 5 obs. at a time on next page. [see **CAMSsequentialT1.R**]


Mean PARS: n = 5
Mean PARS: n = 10
Mean PARS: n = 15
Mean PARS: n = 20
Mean PARS: n = 25
Mean PARS: n = 30
Mean PARS: n = 35
Mean PARS: n = 40
Mean PARS: n = 45
Mean PARS: n = 50
Mean PARS: n = 55
Mean PARS: n = 60
Mean PARS: n = 65
Mean PARS: n = 70
Mean PARS: n = 75
Mean PARS: n = 80
Mean PARS: n = 85
Mean PARS: n = 90
Mean PARS: n = 95
Mean PARS: n = 100
Density
PARS
0.0
0.4
0.8
0
5
10
15
20
25

Comparison of means hypothesis test

Recent work by Mills (2012) has provided an alternative Bayesian testing procedure that does not suffer from the JLB paradox. This allows the use of uninformative priors, so that scientific objectivity can be maintained. The outcome of the test is determined solely by the evidence from the data, with the prior playing little to no role (unless there is convincing reason to support adoption of a more informative prior). The posterior odds for this procedure is the ratio of the posterior density function evaluated at its supremum and at the value under the null hypothesis,

$$Odds = sup_{\theta}[p(\theta|x)]/p(\theta_0|x) \qquad (n).$$

For testing one mean, we can use the analytical form of the marginal posterior, given by equation (2). In this case, under the alternative hypothesis H_1: $\mu_1 = \hat{\mu}_1$, the mode of the posterior density for μ_1, $t^2 = 0$ in equation (2), so that $p(\mu_1 = \hat{\mu}_1|\bar{x}_1) = 1 \times c$, where c is the constant of proportionality. The posterior odds ratio is then simply the value of the equation (2) ignoring c, since the odds ratio is

$$\frac{p(\mu_1 = 0|\bar{x}_1)}{p(\mu_1 = \hat{\mu}_1|\bar{x}_1)} = \frac{c\left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}}{c} = \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}.$$

As can be seen from the above formula, there is a one-to-one correspondence with the p-value in this case [Goodman (1999*a*), Mills (2012)]. This same formula applies for testing the null hypothesis of zero for linear regression coefficients (see Mills (2012), and so has broad application.

Results for PARS: Sert. + CBT vs. PCBO, and Sert. vs. PCBO

The table below gives sequential posterior odds and p-values for difference in means between Sert. + CBT (T1) vs. placebo (T4), and sertraline (T2) vs. placebo (T4), adding 6 obs. for the treatment groups, and 3 for the placebo group each iteration.

	oddsT1T4	pvalueT1T4	oddsT2T4	pvalueT2T4
[1,]	1.364433	1.545970e+00	2.098652	1.74576426
[2,]	1.048564	7.815659e-01	1.541302	1.63338185
[3,]	1.206868	5.777761e-01	1.020097	0.85672767
[4,]	3.459602	1.306617e-01	1.040356	0.79516143
[5,]	3.457469	1.274554e-01	1.020220	0.86395071
[6,]	7.244704	5.260834e-02	1.569954	0.35242323
[7,]	66.921038	4.504293e-03	3.718733	0.11108287
[8,]	313.275480	8.427827e-04	6.373646	0.05873438
[9,]	459.927803	5.939654e-04	3.990254	0.09981578
[10,]	150.087790	1.785411e-03	2.089289	0.23080506
[11,]	1039.340141	2.034694e-04	3.397314	0.12197821

```
[12,]  629.430992 3.597136e-04  4.041063 0.09998239
[13,]  350.535343 7.204952e-04  3.063142 0.13890993
[14,]  322.209483 7.859285e-04  2.965433 0.14521356
[15,]  655.879516 4.525207e-04  2.937452 0.14630118
[16,] 2035.278616 1.258424e-04  3.424589 0.12156739
[17,] 3787.169687 5.272480e-05  6.691671 0.05420558
[18,] 4480.117969 4.185515e-05  8.571422 0.04052134
[19,] 8818.115498 3.040121e-05  8.531238 0.04158536
[20,] 5819.504138 3.143888e-05 16.041600 0.01970928
```

Theoretical details

Suppose $x \sim N(\mu, \tau)$, where the precision. τ, is the inverse of the variance, σ^2, i.e. $\tau = 1/\sigma^2$. We work with the precision rather than the variance simply because it turns out to be mathematically more tractable to do so in what follows, otherwise there is no difference; we can substitute in for the variance at any time and obtain exactly the same result. For a sample of observations, $x = (x_1, x_2, \ldots, x_n)$, the likelihood function is then

$$p(x|\mu, \tau) = (\tau/2\pi)^{\frac{n}{2}} \exp\left[-\frac{\tau \sum_i (x_i - \mu)^2}{2}\right].$$

Omitting the components of the proportionality constant that do not contain the unknown parameters, we have,

$$p(x|\mu, \tau) \propto \tau^{\frac{n}{2}} \exp\left[-\frac{\tau \sum_i (x_i - \mu)^2}{2}\right].$$

A standard uninformative prior that is used extensively in practice is a uniform for each mean, μ, and a log uniform for the precision, so the prior is $p(\mu, \tau) \propto \tau^{-1}$. The log uniform is employed for the precision or variance because $\sigma_j > 0$ for all j, and extremely large values of each σ_j are less likely that values closer to zero, so log scaling is appropriate [see Greenberg (2013), p.46].

Combining prior and likelihood via Bayes' rule, we get a Normal joint posterior,

$$p(\mu, \tau|x) \propto \tau^{\frac{n-2}{2}} \exp\left[-\frac{\tau \sum_i (x_i - \mu)^2}{2}\right].$$

Completing the square inside the exponent and rearranging, this can be written as,

$$p(\mu, \tau|x) \propto \tau^{\frac{n-2}{2}} \exp\left[-\frac{\tau}{2}[\nu s^2 + n(\mu - \bar{x})^2\right],$$

where $\bar{x}$ is the sample mean [see Greenberg, p.47-48, or Hahn, p.62-63].

Integrating with respect to τ leads to a marginal posterior for μ in the form of a Student-t distribution,

$$p(\mu|x) \propto \left(1 + \frac{t^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}, \tag{3}$$

where $t^2 = \frac{n\sum_i (x_i-\mu)^2}{s^2} = \sqrt{n}(\mu - \bar{x})/s$, $s^2 = \frac{\sum_i (x_i-\bar{x})^2}{n}$, $\bar{x} = \sum_i x_i /n$, $\nu = n-1$ are sufficient statistics in this case.

Integrating the joint posterior with respect to μ leads to a marginal posterior for τ in the form of a Gamma distribution,

$$p(\tau|x) \propto \tau^{\frac{\nu}{2}-1} \exp\left(-\frac{\tau\nu s^2}{2}\right),$$

with s^2 and ν sufficient statistics for the distribution in this case.

In summary, given a sample, x, and adopting an uninformative prior, we have a Student-t marginal posterior for μ,

$$\mu \sim t(\bar{x}, s^2, \nu),$$

and a Gamma marginal posterior for τ

$$\tau \sim Gamma(\nu/2, \nu s^2/2),$$

with the summary statistics $\bar{x}$, s^2 and ν sufficient information to completely determine these distributions. Note that the marginal posterior for σ^2 is therefore Inverted Gamma with the same parameters as for τ, i.e. $\sigma^2 \sim IG(\nu/2, \nu s^2/2)$.

Suppose a second sample of m is obtained from the same distribution, $y \sim N(\mu, \tau)$. We can use the posterior distributions obtained above with the first sample as the new prior distribution. Combining with the likelihood for y, it can be shown [below] that the marginal posteriors based on the combined sample $z = (x, y)$ are a Student-t marginal posterior for μ,

$$\mu \sim t(\bar{z}, s_z^2, \nu_z),$$

and a Gamma marginal posterior for τ

$$\tau \sim Gamma(\nu_z/2, \nu_z s_z^2/2),$$

with the summary statistics,

$$\bar{z} = \sum_{i=1}^{n+m} z_i = \frac{\sum_i^n x_i + \sum_i^m y_i}{n+m} = \frac{n\bar{x} + m\bar{y}}{n+m},$$

$$s_z^2 = \frac{\sum_{i=1}^{n+m}(z_i - \bar{z})^2}{n+m} = \frac{ns^2 + ms_y^2}{n+m},$$

$$\nu_z = n + m - 1 = \nu + m,$$

with $s_y^2 = \sum_{i=1}^{m}(y_i - \bar{y})^2 / m$.

The above summary statistics provide sufficient information to completely determine these distributions, so these updating equations can be used as more observations become available and the functional forms of the posteriors remain the same.

A standard alternative prior, which is equally uninformative in practice, particularly for any reasonable sample size, is a Gaussian prior for μ, and a Gamma prior for the precision, τ, with prior parameter values selected to ensure a large variance for each of the prior distributions,

$$p(\mu|\mu_0, \tau_0) = (\tau_0/2\pi)^{\frac{1}{2}} \exp\left[-\frac{\tau(\mu - \mu_0)^2}{2}\right],$$

$$p(\tau|a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \tau^{a_0 - 1} \exp(-b_0\tau).$$

Combining this prior with the normal likelihood, we obtain marginal posteriors of the same form as previously, i.e. Student-t for μ and Gamma for τ, with parameters,

$$\mu_1 = \tau_0\mu_0 + \cdots$$

[* To be completed - see DeGroot and Schervish or Hahn for formulas.]**

Suppose we have m groups consisting of $m-1$ treatment groups and a nontreatment or placebo group, with n_j participants in the jth group. We wish to compare the mean treatment effect between groups to evaluate the efficacy of each treatment. The central limit theorem and the maximum entropy prinviple provide strong justifications for assuming these sample means follow Gaussian processes, so that the likelihood function is,

$$\bar{x}_j \sim N(\mu_j, {\sigma_j}^2), \text{ for group } j,\ j = 1,2,\ldots,m \qquad (1)$$

with μ_j and ${\sigma_j}^2$ unknown parameters.

We wish to test the "no treatment effect" hypothesis, H_0: $\mu_j = \mu$, for all j, as well as various sub-groups of means, such as pairings, e.g., H_0: $\mu_j = \mu_k$, for $j \neq k$. In general, the

variance of the treatment effect for each group is also unknown, which we rewrite in terms of the precision, $\tau = 1/\sigma^2$ for mathematical convenience, so (1) becomes,

$$\bar{x}_j \sim N(\mu_j, \tau_j), \text{ for group } j, \ j = 1,2,\dots,m \tag{2}$$

A standard uninformative prior that is used extensively in practice is a uniform for each mean, μ_j, and a log uniform for each precision (the inverse of the variance), $\tau_j = 1/{\sigma_j}^2$. The log uniform is employed for the precision or variance because $\sigma_j > 0$ for all j, and extremely large values of each σ_j are less likely that values closer to zero, so log scaling is appropriate. A standard alternative prior, which is equally uninformative in practice, particularly for any reasonable sample size, is a Gaussian prior for each mean, μ_j, and a Gamma prior for the precision, τ_j, with prior parameter values selected to ensure a large variance for each of the prior distributions. Applying Bayes' rule using either of the above standard priors and the likelihood given by (1), we obtain a joint Normal-Gamma posterior distribution for the parameters. The marginal posterior distribution for each parameter can then be obtained analytically and is a well-known result, leading to a Student-t distribution for the posterior means. Using the conjugate Normal-Gamma prior leads to essentially the same formulas, with the additional prior parameters having little to no effect if the prior variances are large. The marginal posterior density for each mean is then,

$$p(\mu_j|\bar{x}_j) \propto \left(1 + \frac{t^2}{\nu_j}\right)^{-\frac{(\nu_j+1)}{2}}, \tag{2}$$

where $t^2 = \frac{n\sum_i (x_{ij}-\mu_j)^2}{s_j^2}$, $s_j^2 = \frac{\sum_i (x_{ij}-\bar{x}_j)^2}{n}$, $\nu_j = n_j - 1$.

Note that there is no required assumption that the variances be equal or their ratios across groups known, and the number of participants in each group can differ. The uncertainty due to the unknown variances is captured in the wider tails of the Student-t density in comparison to the Gaussian.

Code:

CAMSsequentialT1.R – produces odds and p-value table, and figure of sequential posteriors.

postoddsmc.R – function to compute posterior odds using MC sample

bayespval.R – function to compute p-values from posterior using MC sample