Set 16 - Comparing means

STAT 401 (Engineering) - Iowa State University

March 1, 2017

One mean

Consider the model $Y_i \stackrel{ind}{\sim} N(\mu, \sigma^2)$. We have discussed a number of statistical procedures to draw inferences about μ:

- pvalue for a hypothesis test, e.g. $H_0 : \mu = \mu_0$,
- confidence interval for μ,
- credible interval for μ,
- posterior model probability, e.g. $p(H_0|y)$, and
- probability statements, e.g. $P(\mu < \mu_0|y)$.

Now, we will consider what happens when you have multiple μs.

Two means

Consider the model

$$Y_{i,j} \sim N(\mu_i, \sigma_i^2)$$

for $i = 1, 2$ and $j = 1, \ldots, n_i$. and you are interested in the relationship between μ_1 and μ_2. We can perform the following statistical procedures:

- pvalue for a hypothesis test, e.g. $H_0 : \mu_1 = \mu_2$,
- confidence interval for $\mu_1 - \mu_2$,
- credible interval for $\mu_1 - \mu_2$,
- posterior model probability, e.g. $p(H_0|y)$, and
- probability statements, e.g. $P(\mu_1 < \mu_2|y)$.

where $y = (y_1, y_2)$.

Data example

Suppose you have two manufacturing processes to produce sensors and you are interested in the average sensitivity of the sensors.

So you run the two processes and record the sensitivity of each sensor in units of mV/V/mm Hg (http://www.ni.com/white-paper/14860/en/). And you have the following summary statistics:

```
d <- readr::read_csv("sensitivity.csv")
sm <- d %>%
  group_by(process) %>%
  summarize(
    n    = n(),
    mean = mean(sensitivity),
    sd   = sd(sensitivity)
  )
sm %>% filter(process <= 2)


# A tibble: 2  4
  process     n     mean       sd
    <int> <int>    <dbl>    <dbl>
1       1    22 7.743761 1.867724
2       2    34 9.207516 3.433806
```

Pvalues and confidence intervals

Because there is no indication that you have any expectation of the sensitivities in process 1 compared to process 2, we will conduct a two-sided two-sample t-test assuming the variances are not equal:

```
t.test(sensitivity ~ process, data = d %>% filter(process <= 2))


Welch Two Sample t-test

data:  sensitivity by process
t = -2.0591, df = 52.747, p-value = 0.04444
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.88977008 -0.03773957
sample estimates:
mean in group 1 mean in group 2
       7.743761        9.207516
```

Posteriors for μ

Recall that

$$\mu_i|y_i \sim t_{n_i-1}(\overline{y}_i, s_i/\sqrt{n_i})$$

and that a draw for μ_i can be obtained by taking

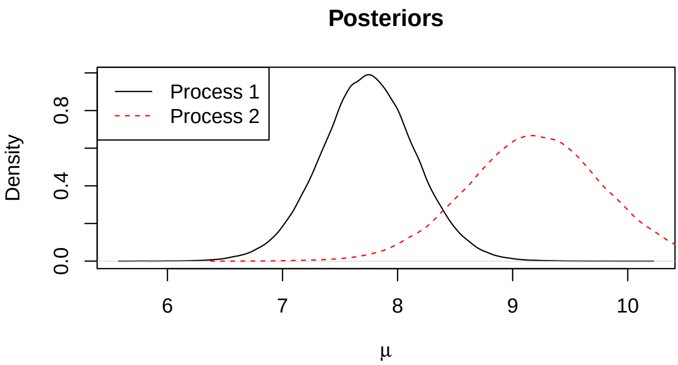
$$\overline{y}_i + T_{n_i-1} s_i/\sqrt{n_i}, \quad T_{n_i-1} \sim t_{n_i-1}(0,1).$$

```
nr = 1e5
mu1 <- sm$mean[1] + rt(nr, df = sm$n[1]) * sm$sd[1] / sqrt(sm$n[1])
mu2 <- sm$mean[2] + rt(nr, df = sm$n[2]) * sm$sd[2] / sqrt(sm$n[2])
```

We can use these draws to compare the posteriors

We can obtain posteriors for μ and plot histograms (or smoothed histograms) to compare the posteriors.

```
plot(density(mu1), main = "Posteriors", xlab=expression(mu)) lines(density(mu2), col='red', lty=2) legend("topl
paste("Process", 1:2), col=c("black","red"), lty=1:2)
```



Credible interval for the difference

To obtain statistical inference on the difference, we use the samples and take the difference

```
diff = mu1 - mu2

# Bayes estimate for the difference
mean(diff)


[1] -1.460891


# Estimated 95% equal-tail credible interval
quantile(diff, c(.025,.975))


       2.5%       97.5%
-2.91288832 -0.01762562


# Estimate of the probability that mu1 is larger than mu2
mean(diff > 0)


[1] 0.02379
```

Multiple means

Now, let's consider the more general problem of

$$Y_i \sim N(\mu_i, \sigma_i^2)$$

for $i = 1, 2, \ldots, J$ and you are interested in the relationship amongst the μ_j.

We can perform the following statistical procedures:

- pvalue for a hypothesis test, e.g. $H_0 : \mu_i = \mu$ for all i,
- confidence interval for $\mu_i - \mu_j$ for a specified i and j,
- credible interval for $\mu_i - \mu_j$ for a specified i and j,
- posterior model probability, e.g. $p(H_0|y)$, and
- probability statements, e.g. $P(\mu_i < \mu_i|y)$ for a specified i and j.

Data example

Suppose you have three manufacturing processes to produce sensors and you are interested in the average sensitivity of the sensors.

So you run the three processes and record the sensitivity of each sensor in units of mV/V/mm Hg (`http://www.ni.com/white-paper/14860/en/`). And you have the following summary statistics:

```
sm
```

```
# A tibble: 3  4
  process     n      mean       sd
    <int> <int>     <dbl>    <dbl>
1       1    22  7.743761 1.867724
2       2    34  9.207516 3.433806
3       3     7 11.815365 4.269126
```

Pvalues and confidence intervals

When there are lots of means, the first null hypothesis is typically

$$H_0 : \mu_i = \mu \,\forall\, i$$

```
m <- lm(sensitivity ~ process, data = d)
anova(m)


Analysis of Variance Table

Response: sensitivity
          Df Sum Sq Mean Sq F value   Pr(>F)
process    1  87.38  87.377  9.2536 0.003463 **
Residuals 61 575.99   9.442
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Then we typically want to look at pairwise differences:

```
pairwise.t.test(d$sensitivity, d$process, pool.sd = FALSE, p.adjust.method = "none")


Pairwise comparisons using t tests with non-pooled SD

data:  d$sensitivity and d$process

  1     2
2 0.044 -
3 0.045 0.169
```

Posteriors for μ

Recall that

$$\mu_i|y_i \sim t_{n_i-1}(\overline{y}_i, s_i/\sqrt{n_i})$$

and that a draw for μ_i can be obtained by taking

$$\overline{y}_i + T_{n_i-1} s_i/\sqrt{n_i}, \quad T_{n_i-1} \sim t_{n_i-1}(0,1).$$

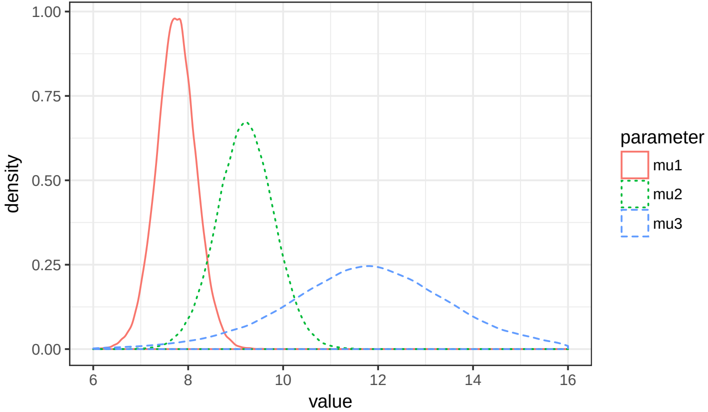
```
posterior_samples <- function(d) {
  data.frame(
    rep = 1:1e5,
    parameter = paste0("mu", d$process),
    value = d$mean + rt(1e5, df = d$n) * d$sd / sqrt(d$n),
    stringsAsFactors = FALSE)
}

draws <- sm %>% group_by(process) %>% do(posterior_samples(.)) %>% ungroup() %>%
  select(-process)
```

We can use these draws to compare the posteriors

We can obtain posteriors for μ and plot histograms (or smoothed histograms) to compare the posteriors.

```
ggplot(draws, aes(x=value, color = parameter, linetype = parameter)) +
  geom_density() + theme_bw() + xlim(6,16)
```



Credible interval for the difference

To compare the means, we compare the samples drawn from the posterior.

```
# Estimate of the comparison means
draws %>%
  tidyr::spread(parameter, value) %>%
  summarize(`P(mu1>mu2|y)` = mean(mu1 > mu2),
            `P(mu1>mu3|y)` = mean(mu1 > mu3),
            `P(mu2>mu3|y)` = mean(mu2 > mu3)) %>%
  gather(comparison, probability)


# A tibble: 3  2
    comparison probability
         <chr>       <dbl>
1 P(mu1>mu2|y)     0.02450
2 P(mu1>mu3|y)     0.02122
3 P(mu2>mu3|y)     0.08523
```