

Solutions to Ex. 1

Have you tried the questions yet? If not, I recommend you make a serious attempt at the questions before looking at these solutions. You can learn a lot by trying and failing in maths. If you read the solution first, you lose that experience.

1.

$$Y_{i1} - Y_{i2} = \mu + b_i + \epsilon_{i1} - \mu - \beta - b_i - \epsilon_{i2}.$$

Thus

$$\begin{aligned} E[Y_{i1} - Y_{i2}] &= -\beta, \\ \text{var}[Y_{i1} - Y_{i2}] &= \text{var}(\epsilon_{i1} - \epsilon_{i2}) \\ &= \text{var}(\epsilon_{i1}) + \text{var}(\epsilon_{i2}) \\ &= 2\sigma^2. \end{aligned}$$

Normality of $Y_{i1} - Y_{i2}$ follows since they are differences of Normal variates, and we have shown that the variance of the difference is constant, independent of i . Thus we have established the required conditions.

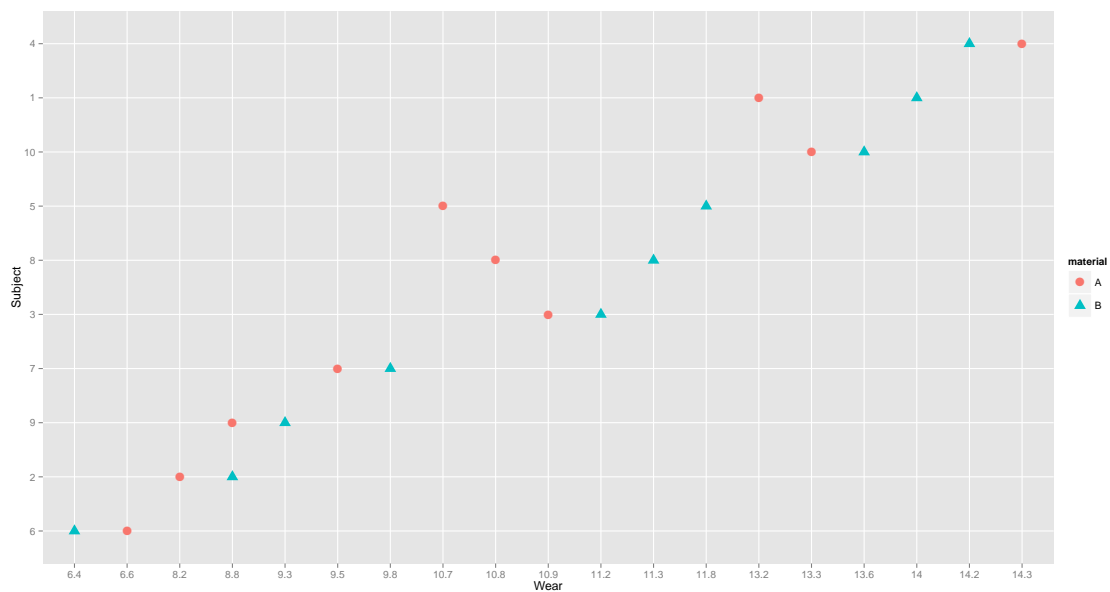
```
2. +++++ Paired t-test +++++
    > attach(BHHshoes)
    > t.test(wear[material=="A"],wear[material=="B"],paired=T)
```

Paired t-test

```
data: wear[material == "A"] and wear[material == "B"]
t = -3.3489, df = 9, p-value = 0.008539
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.6869539 -0.1330461
sample estimates:
mean of the differences
      -0.41
```

The paired t -test has given reasonably strong evidence against the null that the mean difference is 0. The sample mean difference is -0.41, suggesting mean wear is higher in material B.

3. A single plot that will suffice here is plot of wear against subject, using different symbols and colours for the two material types. See the supporting R script for the plotting commands.



We can see that there is large variation between subjects relative to variation within subjects, and that material B has higher wear for most subjects.

4. The model fitting command and output are below. **material** is modelled as a fixed effect, as the interest is in the two particular material types chosen for the study, and **Subject** is modelled as a random effect, as the individuals in the study are not in themselves of interest, and are a sample from a larger population of individuals.

```
> fm1<-lmer(wear~material+(1|Subject), data = BHHshoes)
> summary(fm1)
Linear mixed model fit by REML ['lmerMod']
Formula: wear ~ material + (1 | Subject)
Data: BHHshoes
```

REML criterion at convergence: 54.9

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-1.25097	-0.27027	-0.00901	0.26815	1.26948

Random effects:

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	6.10089	2.4700
Residual		0.07494	0.2738

Number of obs: 20, groups: Subject, 10

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	10.6300	0.7859	13.527
materialB	0.4100	0.1224	3.349

Correlation of Fixed Effects:

	(Intr)
materialB	-0.078

Parameter estimate are $\hat{\mu} = 10.63$, $\hat{\beta}_2 = 0.41$, $\hat{\sigma}_b^2 = 6.1$ and $\hat{\sigma}^2 = 0.075$.

5. The estimate of β_2 is 0.41, which is the difference between the sample means for the two materials (given in the t -test output). Calculating the sample mean for material A

```
> mean(wear[material=="A"])
[1] 10.63
```

we see that the estimate of μ is the sample mean for material A. Now, noting that $Var(Y_{ij}) = \sigma_b^2 + \sigma^2$,

$$\begin{aligned}\hat{\mu} &= \frac{1}{10} \sum_{i=1}^{10} Y_{ij}, \\ \Rightarrow Var(\hat{\mu}) &= \frac{\sigma_b^2 + \sigma^2}{10},\end{aligned}$$

and so the estimated standard error is

$$\sqrt{\frac{6.10089 + 0.07494}{10}} = 0.7859,$$

to 4 d.p. Also,

$$\begin{aligned}\hat{\beta}_2 &= \frac{1}{10} \left(\sum_{i=1}^{10} Y_{i2} - \sum_{i=1}^{10} Y_{i1} \right), \\ &= \frac{1}{10} \left(\sum_{i=1}^{10} (\mu + \beta_2 + b_i + \epsilon_{i2}) - \sum_{i=1}^{10} (\mu + b_i + \epsilon_{i1}) \right), \\ \Rightarrow Var(\hat{\beta}_2) &= \frac{2\sigma^2}{10},\end{aligned}$$

and so the estimated standard error is

$$\sqrt{\frac{2 \times 0.07494}{10}} = 0.1224,$$

to 4 d.p.

6. The `t value` is the same as the (absolute value of the) t -statistic in the paired t -test, as in both cases they are difference in sample means divided by the estimated standard error for the difference. The analysis with the mixed effects model (regarding the effect of `material`) is equivalent to the paired t -test. The mixed effects model does, however, give a little more information, as it also give estimates of between and within subject variability.