Chapter 1

Section 1.3.8 Mitochondrial Variation. This section claims (last sentence) that because there is no recombination among mitochondrial DNA of different individuals, that localization of genetic effects is difficult. This is not a complete picture, because it turns out that many mtSNPs can be shown to have multiple points of origins rather than a single origin as in the "infinite sites" model. This is due to a mutation rate, due to copying errors, which is approximately 1000 times higher than that of most nuclear DNA [2]. This should also have been mentioned in Chapter 2.

Chapter 2

Page 40, the 2nd variance term in the last equation should read $Var_p(2p_t)$.

Page 42, on the 2nd line of the 3rd paragraph, the sentence should read:

… and off diagonal terms, $k_{ij} = 2F_{ij} = 2F(1 + 2\alpha_i\alpha_j - \alpha_i - \alpha_j)$

Page 44, the first line under the formula, it should read “$z_1$ and $z_2$”.

Figure 2.3 Caption incomplete, it should read: Allele frequency distribution under selection or not. Wright's formula is used to depict the cumulative density of allele frequencies between frequency 0.001 and 0.999 under no selection (solid line) and selection with $\sigma = 12$ (dotted line).

Section 2.9.1 Page 69. The function plotPC in the R-code snippet is wrong please use the following function definition

```r
plotPC<-function(e,GRP="ASW",rest=c("CEU","CHB","CHD","GIH","JPT","LWK","MEX","MKK","TSI","YRI"),COLOR="red",restc=c("blue","green","yellow","orange","black","purple","pink","gray","violet","cyan","blue"),d1=1,d2=2,
xl=c(-0.05,0.05),yl=c(-0.05,0.05),
gp=grp)
{
plot(e$vectors[gp==GRP,d1],e$vectors[gp==GRP,d2],xlim=xl,ylim=yl,col=COLOR,main="",xlab=paste("PC",d1,sep=""),ylab=paste("PC",d2,sep=""))
lr<-length(rest)
if (length(rest) >= 1) {
```


cat(paste("Plot of ", grp[1], " PC2 vs PC1 
"))
for (i in 2:length(rest)) {
  resp<-readline (prompt= paste("Add ",rest[i-1],"? y/n "))
  if(resp=="y" | resp=="Y") lines(e$ vectors[gp==rest[i-1],d1],e$ vectors[gp==rest[i-1],d2],xlim=xl,ylim=yl,col=restc[i-1],type="p")
}
Page 153, Figure 4.4. The legend refers to extraneous Fig. 4.4c and 4.4d which are not shown (and not discussed in the text)

Page 156, on the 5th line of the 2nd paragraph, move “for \( i \neq j = 1,\ldots, N \)” before the “and” at the end of the previous line.

Page 158, in the denominator of equation 4.16, the first \( k \) should be a subscript

Page 163, the 2nd line under equation 4.20 should read \((\hat{\beta}_j)^2\)

Page 168, homework problem 1. The formula for \( Y_{\text{adjusted}} \) should be

\[
Y_{\text{adjusted}} = Y - V_1'Y - V_2'Y - \ldots - V_L'Y
\]

Page 168, homework problem 3. The model should be

\[
Y = \sum_{k=1}^{r} \beta_k X_k + \sum_{k=1}^{N} \alpha_k V_k
\]

Chapter 5

Page 191 7 lines from top: “but with sibling data also available” should read “but with offspring data also available”

Chapter 6

Page 218 9 lines from top, the expression for \( \Pr(Q | \theta) \) is better written as

\[
\Pr(Q | \theta) = \pi_{q_1} \left[ \prod_{i=2}^{T} a_{q_{i-1}q_i} \right]
\]

Page 219. Equation (6.5) replace \( \theta \) with \( \lambda \) at end of equation

Page 220. 6th line of section 6.4.1 replace \( \theta \) with \( \lambda \) for expression for \( \alpha_i(i) \)

Page 221 line before equation 6.9 the "\( \gamma_i\)" should be "\( \gamma_i(i) \)". Same typo appears on page 234 line 10.

Page 221 second line after equation 6.10 the "\( \Pr(Q|O,t)\)", should be "\( \Pr(Q|O,\lambda)\)"

Chapter 7

Page 245, 8th line from top should read: “the non-centrality parameter is \((\beta - \beta_0)^2 / \text{Var} (\hat{\beta})\)”
Page 245, equation (7.1) should read \[ \Phi^{-1}(1 - \alpha / 2) + \Phi^{-1}(s)^2 \]

Page 254, weights in the R code should be

```r
> wgts=c(217.6,281.6,140.8,108.8,140.8,70.4,13.6,17.6,8.8, 192,
  245.6, 245.6, 96, 98.2, 98.2, 12, 6.2, 6.2)
```

Page 256, Table 7.5 and surrounding text. In this discussion I leaped to the wrong conclusion. In fact for all the models in Longmate 2001 the expected likelihood method (used by Quanto) is equivalent to the exemplary data method. This will be true for any model in which the log likelihood is linear in the cell counts (as provided in Table 7.5), and such linearity holds for the logistic regression model. The fact that there appeared to be a disagreement between the expected likelihood approach and the exemplary data approach when considering power for the interaction term (SxG) is due to numerical differences. The R-code using the function `optim` is not converging close enough to the optimal values (which are found easily by the glm procedure) improving the starting values and increasing the number of iterations for `optim` yields an identical NCP.

Page 258, line 6 from top, the “2.80 (power=23%)” should be “3.50 (power=28.2%)” (but see note above)

Page 259, line 7 from top “(power=28%)” should be “(power=28.6%)” (but see note above)

Page 278, 13 lines below equation 7.12, replace \( \beta_{2j} = 2 \) with \( \beta_{2j} = 0.2 \).
Figure 7.4 is missing the curve for the $n1/(n1+n2) = 1/2$ design. Also the non-monotonic curve should be dotted and not solid. A revised version is given below.

![Power of a 2-stage study](image)

**Fig. 7.4** Power of a two stage study to detect an odds ratio of 1.25 for a 20% allele, according to the fraction of subjects genotyped in stage 1 and the first stage type I error level. Odds ratio is assumed to be equal to 1.25, and frequency 0.2; 5,000 cases and 5,000 controls genotyped. Also shown (dotted line) is the power of a “discovery + replication” design with $n1/(n1 + n2) = 1/2$

Page 256, Table 7.5 has suffered an accident: apparently it collided with Table 7.4 (or they are attempting to mate). Ignore all lines of Table 7.5 beyond the first 6.
Page 260, Figure 7.2. The legend mixes up the solid with the dotted lines. It should read

Fig. 7.2 Sample sizes plotted against odds ratios. Sample sizes computed in QUANTO are those required to maintain power of 80 % for detecting the effects of variants with specified ORs and allele frequencies from 10 to 30 % for sibling (dotted lines) versus random (solid lines) controls.
Chapter 8

Page 300, after the first occurrence of ref #32, fix the italicized “other references”

Section 8.1.1. The statement on page 305: "so that the total polygene variance
\[ \text{Var}(v_i) = \sum_k 2p_k(1-p_k)w_k^2 \] equals 8" should say "equals 1".

Page 309, second to last line, fix “(refs)”

Page 312, in the caption of figure 8.1, the expression “4Ne=12” should read “4Ne_s = 12”.

Page 318, expression (8.31) the expression on the RHS of the equal sign needs to be multiplied by \( \frac{1}{2} \).

Page 323, homework problem 3. The expression for \( \hat{\alpha} - \alpha \) should read
\[
\hat{\alpha} - \alpha = i^{-1}(\alpha)U(\alpha; Y)
\]