

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2006, The Johns Hopkins University and Karl Broman. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.

## Statistics for laboratory scientists

### Solutions for the homework problems for lecture 2

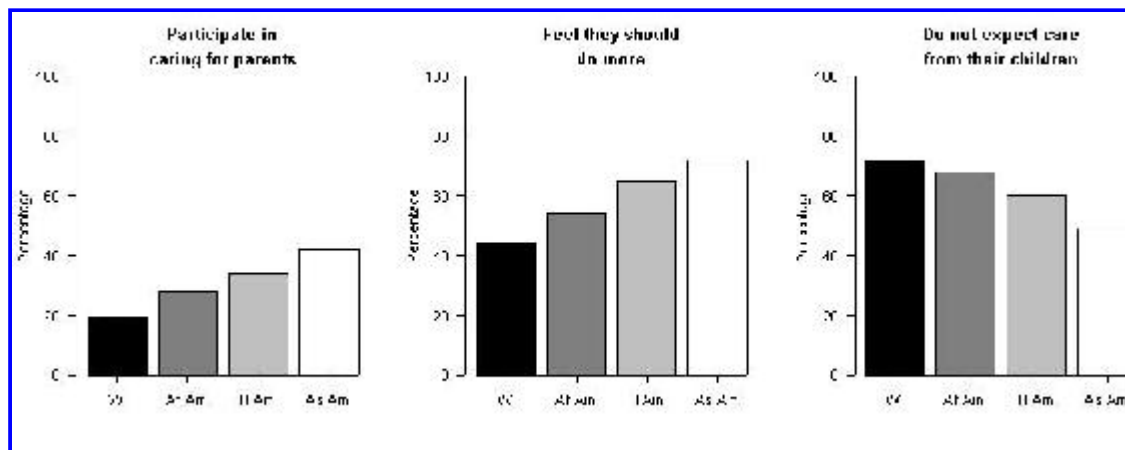
1. Regarding the pie charts:

a. *Name three things wrong with these figures.*

- i. The values in the pie charts don't add up to 100%.
- ii. The three figures are not consistent in their assignments of colors/shading to groups.
- iii. Pie charts are always awful.

b. *Sketch an improved version of these figures.*

I like the following bar charts.



[Click on the figure for a larger version.]

Here is the R code I used to make them:

```
datA <- c(19,28,34,42)
datB <- c(44,54,65,72)
datC <- c(72,68,60,49)

par(mfrow=c(1,3),las=1)
barplot(datA, col=c("black","gray50","gray75","white"),
        names=c("W","Af Am", "H Am", "As Am"), ylim=c(0,100),
        main="Participate in\n\ncaring for parents",
```

```

      ylab="Percentage")
barplot(datB, col=c("black", "gray50", "gray75", "white"),
        names=c("W", "Af Am", "H Am", "As Am"), ylim=c(0,100),
        main="Feel they should\n\nndo more",
        ylab="Percentage")
barplot(datC, col=c("black", "gray50", "gray75", "white"),
        names=c("W", "Af Am", "H Am", "As Am"), ylim=c(0,100),
        main="Do not expect care\n\nfrom their children",
        ylab="Percentage")

```

Note: The `\n` within the character string for the argument `main` (for placing the title on the plot) indicates the insertion of a line break. (I inserted two line breaks to give a better space between the lines in the title.) The argument `las=1` to the function `par` causes the numbers on the y-axis to be printed up-and-down rather than sideways. The rest you can probably figure out. If you can't, check out the R help files.

2.
  - a. mean=10.1; median=10.0; SD=1.1
  - b. Add 2 to the mean and median; the SD is unchanged
  - c. Multiply the mean, median and SD by 10
  - d. Multiply the mean and median by -10; multiple the SD by 10
  - e. For the mean and median, add 2 and then multiply by 10 (new mean=121; new median=120). For the SD, just multiply by 10 (new SD=11)
  - f. For the mean and median, multiply by 10 and then add 2 (new mean=103; new median=102). For the SD, just multiply by 10 (new SD=11)
3. Because the distribution has a long right tail, the mean is larger than the median, and so the red and blue segments are the median and mean, respectively.

The SD is about 15. (Think "typical deviation from the average".)

4. The SD is about 3. (Think "typical deviation from the average".)

Windows ]

EST 2006