Is Human Height Bimodal?

Mark F. SCHILLING, Ann E. WATKINS, and William WATKINS

The combined distribution of heights of men and women has become the canonical illustration of bimodality when teaching introductory statistics. But is this example appropriate? This article investigates the conditions under which a mixture of two normal distributions is bimodal. A simple justification is presented that a mixture of equally weighted normal distributions with common standard deviation σ is bimodal if and only if the difference between the means of the distributions is greater than 2σ. More generally, a mixture of two normal distributions with similar variability cannot be bimodal unless their means differ by more than approximately the sum of their standard deviations. Examination of national survey data on young adults shows that the separation between the distributions of men's and women's heights is not wide enough to produce bimodality. We suggest reasons why histograms of height nevertheless often appear bimodal.

KEY WORDS: Bimodal distribution; Living histogram; Normal distribution.

1. THE DISTRIBUTION OF HUMAN HEIGHT

Brian Joiner's (1975) living histogram (Figure 1) of his students at Penn State grouped by height inspired the standard classroom example of bimodality resulting from a mixture of two populations.

Although the separate distributions of his male and female students are approximately normal, the histogram of men and women together is clearly bimodal. Joiner wrote, "Note that this histogram has a bi-modal shape due to the mixing of two separate groups, males and females." This appealing idea appears in several introductory textbooks:

> "A histogram of the heights of students in a statistics class would be bimodal, for example, when the class contains a mix of men and women." (Iversen and Gergen 1997, p. 132)
>
> "Bimodality often occurs when data consists of observations made on two different kinds of individuals or objects. For example, a histogram of heights of college students would show one peak at a typical male height of roughly 70″ and another at a typical female height of about 65″." (Devore and Peck 1997, p. 43)
>
> "If you look at the heights of people without separating out males and females, you get a bimodal distribution . . . " (Wild and Seber 2000, p. 59)

Still other textbooks include a problem that asks students to predict the shape of the height distribution of a group of students when there are an equal number of males and females. The expected answer is "bimodal." The distribution of height offers a plausible example of bimodality and is easy for students to visualize.

Rather than sending our students out onto the football field *à la* Joiner to demonstrate bimodality, we decided to get some

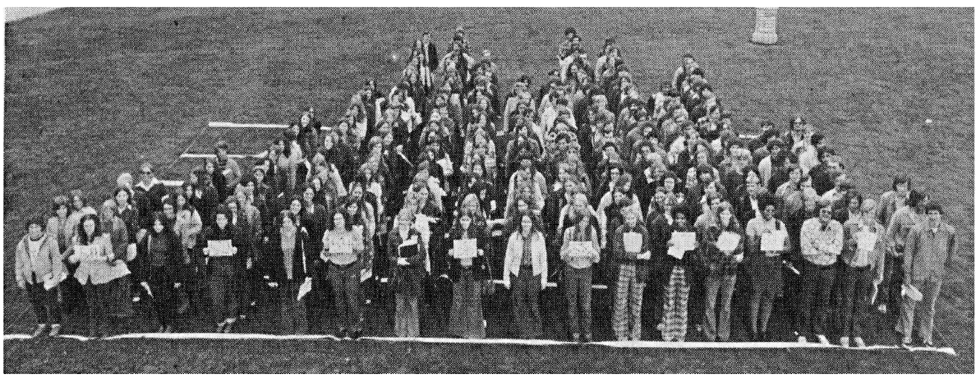

Figure 1. Joiner's living histogram of student height.

government data and construct the approximate theoretical density function for the mixture of the male and female populations. The most recent National Health and Nutrition Examination Survey (NHANES III), conducted in 1988–1994 by the United States National Center for Health Statistics, reports the cumulative distribution of height in inches for males and for fe-

Mark F. Schilling is Professor, Ann E. Watkins is Professor, and William Watkins is Professor, Department of Mathematics, California State University, Northridge Northridge, CA 91330-8313 (E-mail: ann.watkins@csun.edu). The authors thank Rebecca Walker and Deborah Nolan for helpful comments.

 DOI: 10.1198/00031300265

males in the 20–29 age bracket (U. S. Census Bureau 1999). The data for each sex have the means and standard deviations in Table 1 and each follow a normal distribution reasonably well.

Table 1. Summary Statistics for NHANES Height Data

	Mean	*SD**
Males	69.3	2.92
Females	64.1	2.75

*Standard deviations were computed from the cumulative distributions, as they were not supplied by NHANES.

Using the NHANES means and standard deviations as parameters for two normal densities and assuming equal numbers of each sex, we get the graph in Figure 2 for the theoretical mixture distribution of height for persons aged 20–29.

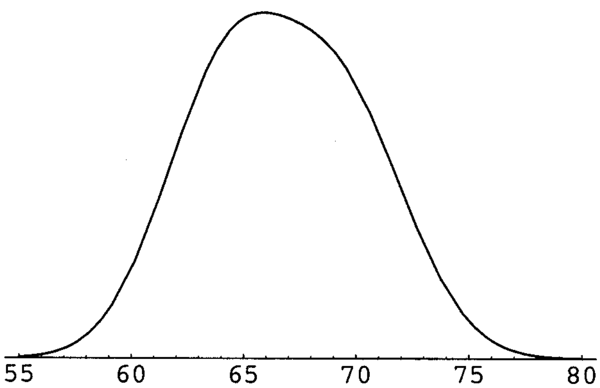

Figure 2. Theoretical distribution of U.S. young adult human height in inches as a mixture of two normal distributions using means and standard deviations from NHANES data.

This obviously is not a bimodal distribution! Yet Joiner's living histogram shows two clear peaks. We start our investigation of this paradox by studying the modality of a mixture of two unimodal densities, with particular focus on normal densities.

2. AN INVESTIGATION OF BIMODALITY

Most students are willing to believe that the mixture of two unimodal densities with differing modes will necessarily be bimodal, as each of the component modes will generate a peak in the mixture distribution. It is quite easy to show that this is not the case. Consider a mixture of two triangular distributions, as shown in Figure 3. If neither distribution's support overlaps the

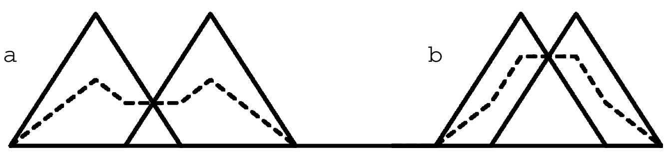

Figure 3. Dashed lines represent an equal mixture of the component distributions (solid lines).

other's mode (Figure 3a), then the mixture distribution is indeed bimodal. However, if each mode is contained within the support of the other distribution (Figure 3b), then outside the component modes the mixture is monotone, while between modes the mixture is the sum of two linear functions, hence is itself linear. Thus the mixture is not bimodal.

Now consider mixture densities of the form $f(x) = pf_1(x) + (1-p)f_2(x)$, where f_1 and f_2 are each normal densities with means $\mu_1 < \mu_2$ and variances σ_1^2 and σ_2^2, respectively, and $0 < p < 1$. Consider first the simplest situation where $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and $p = 0.5$. If μ_1 and μ_2 are far apart, then f clearly will be bimodal, resembling two bell curves side by side. This certainly happens when $\mu_2 - \mu_1$ is greater than about 6σ. Furthermore, it is easy to see that the modes of f occur not at μ_1 and μ_2, but between them: Both f_1 and f_2 are strictly increasing below μ_1, so f has positive derivative there. At μ_1, f_1 has a derivative of zero and f_2 is increasing, so f has positive derivative at μ_1. Thus any mode must be larger than μ_1. Similarly, any mode must be smaller than μ_2.

But what happens when μ_1 and μ_2 are much closer together? The result for this case is generally credited to Cohen's (1956) problem in the *American Mathematical Monthly*, but dates back at least to Helguerro (1904):

Theorem 1. (Helguerro 1904). Let f_1 and f_2 be normal densities with respective means μ_1 and μ_2 and common variance σ^2, and let f be the mixture density $0.5\, f_1 + 0.5\, f_2$. Then f is unimodal if and only if $|\mu_2 - \mu_1| \leq 2\sigma$.

Helguerro and the *Monthly* include a proof of the seemingly obvious fact that a mixture of two normal densities must be either unimodal or bimodal. With that assumption, a simpler justification of Theorem 1 appears below.

Proof: A normal density is concave down between its inflection points $\mu \pm \sigma$ and concave up elsewhere. Note that f is symmetric around $m = (\mu_1 + \mu_2)/2$. Now if $|\mu_2 - \mu_1| > 2\sigma$, then both f_1 and f_2 are concave up at m, hence so is f. Therefore f has a local minimum at m, which implies that f is bimodal. Conversely, if $|\mu_2 - \mu_1| < 2\sigma$, then both f_1 and f_2 are concave down at m, hence so is f. Thus, f has a local maximum at m and is therefore unimodal. In the borderline case $|\mu_2 - \mu_1| = 2\sigma$ the second derivative of f vanishes, but the fourth derivative can be used to show that f has a maximum at m and thus is unimodal.

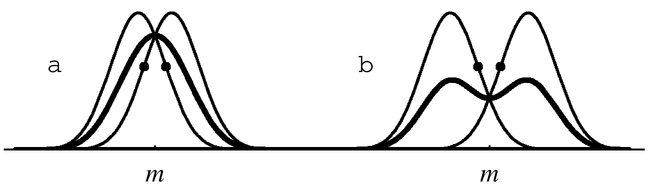

Figure 4. (a) When $|\mu_1 - \mu_2| < 2\sigma$, both normal components (light lines) are concave down at $m = (\mu_1 + \mu_2)/2$, so the mixture f (heavy line) is also concave down at m and is therefore unimodal. (b) When $|\mu_1 - \mu_2| > 2\sigma$, both components are concave up at m, so the mixture f is also concave up at m, hence is bimodal. The dots indicate in˚ection points of the component densities.

Figure 4 illustrates this argument that a two standard deviation separation between the means is needed for bimodality. The symmetry in Figure 4 is crucial and so this proof does not

Table 2. The Mixture Density $f = pf_1 + (1-p)f_2$ Is Bimodal If and Only If $|\mu_2 - \mu_1|$ Exceeds $(\sigma_1 + \sigma_2)$ Times the Value Indicated

σ_1/σ_2	p = 0.3	p = 0.4	p = 0.5	p = 0.6	p = 0.7
1.00	1.36	1.22	1.00	1.22	1.36
0.95	1.35	1.20	1.08	1.24	1.36
0.9	1.34	1.16	1.11	1.25	1.36
0.8	1.29	1.01	1.15	1.26	1.35
0.7	1.18	1.06	1.16	1.24	1.32
0.6	0.96	1.06	1.13	1.20	1.26
0.5	0.96	1.03	1.08	1.14	1.19

generalize to the case when $\sigma_1^2 \neq \sigma_2^2$, as with the heights of men and women, or when there is an unequal mixture of f_1 and f_2, which will occur more often than not in classroom experiments.

The modality of an arbitrary mixture of two normal densities involves a complex interplay between the difference in means, the ratio of the variances, and the mixture proportions. Strong conditions for unimodality/bimodality were derived by Eisenberger (1964) and Behboodian (1970). Robertson and Fryer (1969; reproduced by Titterington, Smith, and Makov 1985) obtained precise specifications of when bimodality occurs. (See also Kakiuchi (1981) and Kemperman (1991), who addressed the question without the normality assumption.) We have attained bimodality conditions equivalent to those of Robertson and Fryer, although we derive and present them quite differently. These results follow from the observation that in order for a mixture $f(x)$ of normal densities to be bimodal there must be at least one value x where $f(x)$ satisfies both $f'(x) = 0$ and $f''(x) > 0$, so that $f(x)$ possesses a relative minimum. Solving $pf_1'(x) + (1-p)f_2'(x) = 0$ and $pf_1''(x) + (1-p)f_2''(x) > 0$ simultaneously yields the inequality $f_2'(x)/f_1'(x) < f_2''(x)/f_1''(x)$. After extensive algebraic manipulations of this inequality, we obtain the following generalization of Theorem 1.

Let $r = \sigma_1^2/\sigma_2^2$ and define the *separation factor* $S(r)$ to be

$$S(r) = \frac{\sqrt{-2 + 3r + 3r^2 - 2r^3 + 2(1 - r + r^2)^{\frac{3}{2}}}}{\sqrt{r}(1 + \sqrt{r})}.$$

Then the mixture density $f = pf_1 + (1-p)f_2$ is unimodal for all p if and only if $|\mu_2 - \mu_1| \leq S(r)(\sigma_1 + \sigma_2)$. Note that $S(1) = 1$, which for the case $p = 0.5$ yields Theorem 1.

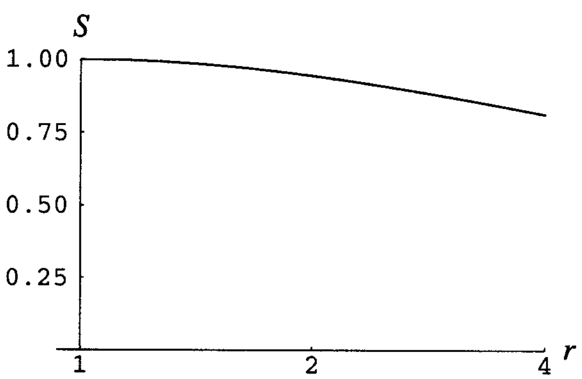

Figure 5. Separation factor $S(r)$ needed for possible bimodality of a mixture of two normal densities, where r is the ratio of their variances. If $|\mu_2 - \mu_1| \leq S(r)(\sigma_1 + \sigma_2)$, any mixture proportion results in unimodality.

Figure 5 shows that $S(r)$ is close to 1 for cases in which the component variances are at all similar (as with heights), thus giving us the following rule of thumb:

Regardless of the mixture proportion p, a mixture of two normal densities with approximately equal variances cannot be bimodal if the separation in means is much less than the sum of the component standard deviations.

Although Figure 5 indicates that bimodality occurs for certain mixture proportions if the separation in means is less than or equal to $\sigma_1 + \sigma_2$, for most mixture proportions a somewhat *greater* separation is required. Table 2 gives the factor for $\sigma_1 + \sigma_2$ that separates unimodal and bimodal mixtures of normal densities for specific values of the mixture proportion p and various choices of σ_1/σ_2. For example, if $\sigma_1/\sigma_2 = 0.9$ and $p = 0.6$, then the mixture is bimodal if the means are farther apart than $1.25(\sigma_1 + \sigma_2)$; if not, it is unimodal.

Mixtures that are bimodal when the separation in means is less than the sum of the standard deviations have only a slight dip if the standard deviations are at all similar. Figure 6 shows the mixture with the most pronounced dip possible among all mixtures of two normal densities in which $0.6 \leq \sigma_1/\sigma_2 \leq 1$.

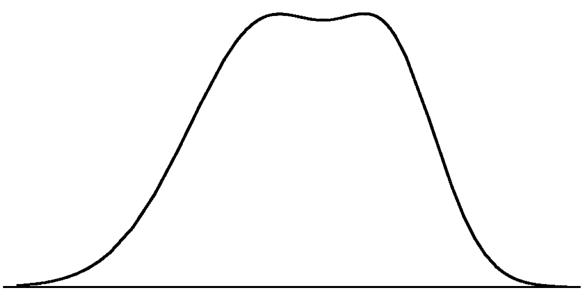

Figure 6. Mixture of two normal densities with $\sigma_1 = 0.6\sigma_2$, $|\mu_2 - \mu_1| = \sigma_1 + \sigma_2$, $p = .288$. This density has the most pronounced dip of any mixture in which $0.6 \leq \sigma_1/\sigma_2 \leq 1$ and $|\mu_2 - \mu_1| \leq \sigma_1 + \sigma_2$.

3. THEORY VS. DATA

Although the theoretical and numerical results of Section 2 give some guidelines for when a mixture density might be unimodal or bimodal, the modality issue becomes more complex when one is dealing with actual data such as students' heights. The study of bimodality and multimodality for data has a long and extensive history, beginning with Pearson (1894). Many investigators have developed methods for resolving the distribution into its underlying (typically Gaussian) components, as well as testing whether the number of these components is greater than one. See Holgersson and Jorner (1978), Everitt and Hand (1981), and Titterington, Smith, and Makov (1985) for good overviews of the subject and extensive references.

This section focuses on the modality of human height as it appears in sample data. In particular, why is Joiner's living histogram bimodal? If we look at the NHANES data of female and male heights, $\sigma_1/\sigma_2 = 2.75/2.92 = 0.94$. According to the second row of Table 2, a separation in means of at least $1.08(2.75 + 2.92) = 6.1$ inches is necessary for bimodality of the corresponding mixture distribution for most typical class-

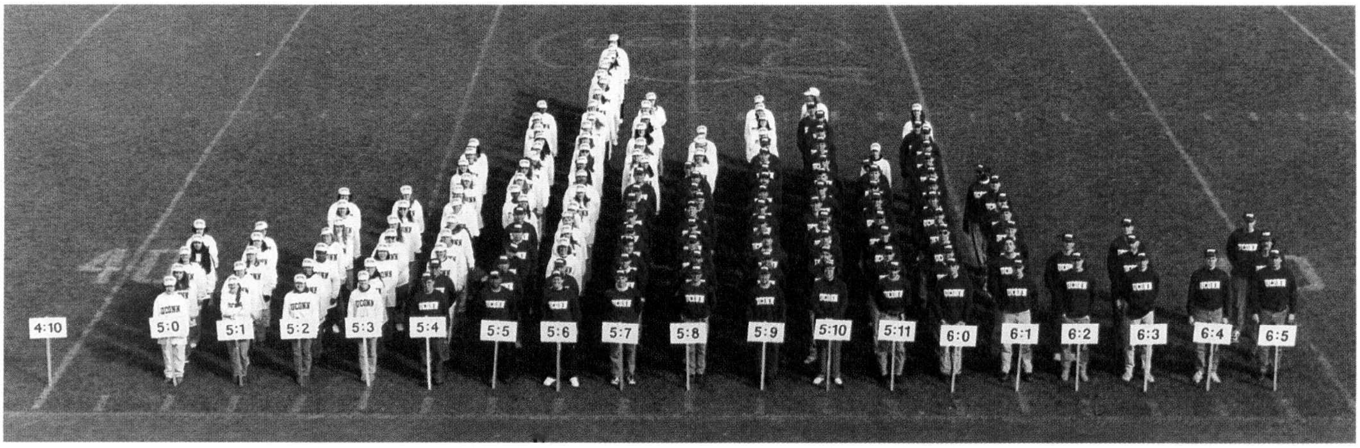

Figure 7. Living histogram of 143 student heights at University of Connecticut.

room proportions of males and females. But the difference of 5.6 inches in the men's and women's means in Joiner's sample (Figure 1) is less than that.

Yet Joiner is not the only professor to get a bimodal distribution when male and female heights are combined. The living histogram in Figure 7, which appeared in *The Hartford Courant* (1996), Crow (1997), and Wild and Seber (2000), shows the male and female genetics students of Linda Strausbaugh at the University of Connecticut in 1996. Crow writes, "Since both sexes are included, the distribution is bimodal." The effect is somewhat bimodal, but only slightly. The StatLab (Hodges, Krech, and Crutchfield 1975) data of 1296 male and 1296 female heights in the San Francisco Bay area also appear somewhat bimodal. We now consider possible reasons for this discrepancy between data and theory.

3.1 Bad Parameters

We cannot say for sure that the NHANES means and standard deviations in Table 1 apply precisely to college students. The NHANES sample consists only of around 600 of each sex aged 20–29 and the heights are self-reported. College students are not demographically identical to the general population aged 20–29. And Joiner's histogram was made 20 years before the NHANES study.

However, theoretical mixtures that use the sample means, standard deviations, and mixture proportions from the Joiner and Strausbaugh samples as parameters are still not bimodal. For Joiner's photograph, we found means of 64.6 inches for females and 70.2 inches for males, and respective standard deviations of 2.6 and 2.8 inches. The group is 54% female. (We are uncertain about the exact counts for some heights but not enough to affect the analysis significantly.) The difference in means is 5.6 inches, which slightly exceeds the sum of the sd's, 5.4 inches, but the separation needed for bimodality in this case is about $1.12(\sigma_1 + \sigma_2) = 6.05$ inches.

The computed means and standard deviations in inches for Strausbaugh's students are 64.8 and 2.7 for females and 70.1 and 3.0 for males; the mixture is about 45% female. Here the difference in means is not even as large as the sum of the sd's, and is again too small for bimodality. Figure 8 shows a mixture of two normal distributions with the same means and standard deviations as the Strausbaugh sample and the same mixture proportion—it could easily pass for a bell curve, although the top is slightly less peaked.

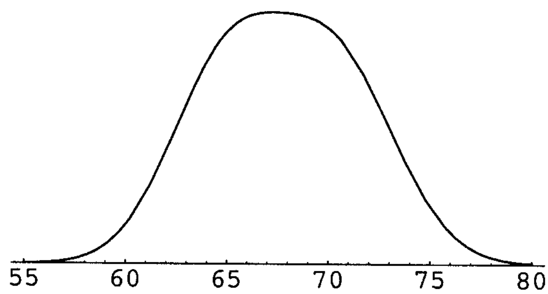

Figure 8. Mixture of two normal densities with the same means, standard deviations, and mixture proportions as the Strausbaugh sample.

Indeed, unimodality is quite robust: as the means of the component normal densities separate, any bimodality remains quite unremarkable at first (see Figure 9). For an equal mixture of two normal densities having the same standard deviation σ to have a dip of 36%, like that in the Joiner histogram, requires a separation in means of about 3σ.

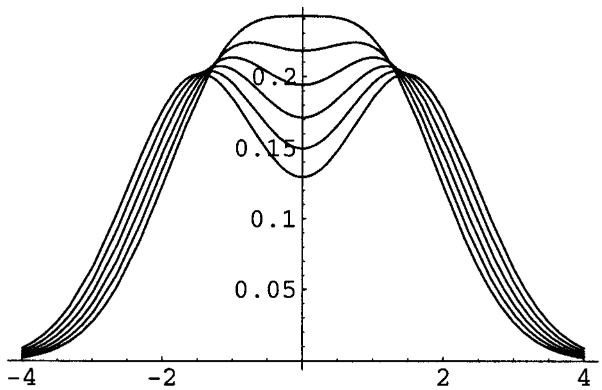

Figure 9. Six equal mixtures (p = .5) of two normal densities. Each component has a standard deviation of 1. The values of $|\mu_2 - \mu_1|$ are 2.0, 2.2, 2.4, 2.6, 2.8, and 3.0.

3.2 Bad Model

The normal distribution is only an approximation to the distribution of human height. What if the true distribution within each sex is somewhat different? Kemperman (1991) provided necessary and sufficient conditions for a mixture of unimodal component densities satisfying certain weak prerequisites to be itself unimodal. These constraints are rather complex; however the concavity arguments of Theorem 1 lead easily to a simple condition for the possibility of a bimodal mixture when normality is not assumed:

Theorem 2. Let f_1 and f_2 be continuous and differentiable densities on the real line. Let f_1 be decreasing and concave up for $x > P$, and let f_2 be increasing and concave up for $x < Q$, where $P < Q$. Then there exist mixtures $f(x) = pf_1(x)+(1-p)f_2(x)$ that are not unimodal.

Proof: Let x be any point in the interval (P, Q). Since at x, f_1 is decreasing and f_2 is increasing, it is possible to choose p so that $f'(x) = pf_1'(x) + (1-p)f_2'(x) = 0$. In addition, f is concave up in (P, Q) since f_1 and f_2 are. Thus f has a local minimum at x, hence is not unimodal.

For example, it is known that there are somewhat more very tall and very short people than what would be found for a perfect normal distribution, and so a properly scaled t distribution might be a reasonable alternative model. Student's t distribution with ν degrees of freedom has inflection points at $\pm\sqrt{2\nu/(2\nu+1)}$. An equal mixture of two t distributions shifted apart by more than $2\sqrt{2\nu/(2\nu+1)}$ therefore satisfies the conditions of Theorem 2 with its local minimum occurring at the center of symmetry of f, and the mixture is bimodal. For the t distribution with $\nu = 10$, the required separation is $2\sqrt{2\cdot 10/(2\cdot 10+1)} = 1.95$. This distribution has a standard deviation of $\sqrt{\nu/(\nu-2)} = 1.12$. The distributions of the heights of males and of females have standard deviations of approximately 2.8 inches. Thus the means need to be separated by $(1.95)(2.8/1.12) = 4.9$ inches, less than what is required with the normal model. The mean difference in the NHANES data is 5.2 inches, so with the t model, bimodality would result, though the dip would be very small. The available percentiles from NHANES, however, actually fit the normal model better than a t model with small to moderate degrees of freedom.

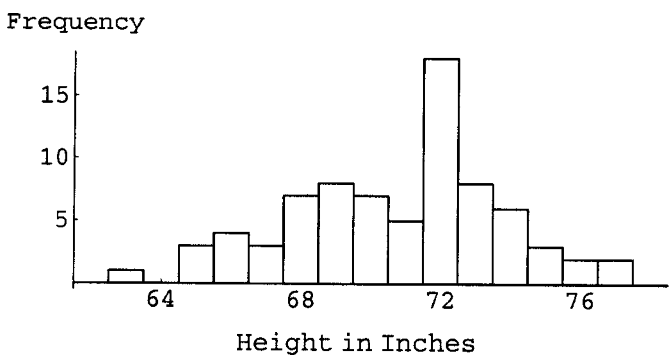

Figure 10. Self-reported heights of 77 male students at U.C. Davis.

3.3 Bad Data

We cannot be confident that the students lined up at their actual heights. For example, Figure 10 shows the heights of the men in a class at the University of California at Davis (Saville and Wood 1996). The instructor sent a sheet of paper around the class and asked students to record their height and gender. The spike at six feet cannot reasonably be attributed to chance variation, and it looks very much like students an inch or two shorter rounded their heights up to six feet.

Separating the University of Connecticut students by gender shows that the impression of bimodality is largely due to a spike in women's heights at $5'\,6''$ and spikes in men's heights at $5'\,10''$ and six feet, apparently popular heights when lining up with classmates. The Joiner sample also has a modest peak at six feet.

More subtly, students must round their height to inches. If men like to be thought of as taller and women prefer to be thought of as shorter (more likely for the women in 1975 than today), then it would be reasonable for the men to round their heights up to the next inch and for the women to round down to the next inch. That would produce an extra $1''$ separation between the means compared to rounding to the nearest inch. The resulting mixture is bimodal, although the dip is so slight (just 0.16% below the lower relative maximum) as to be nearly undetectable. Figure 11 is the counterpart of Figure 2 but with an extra one-inch separation in the means.

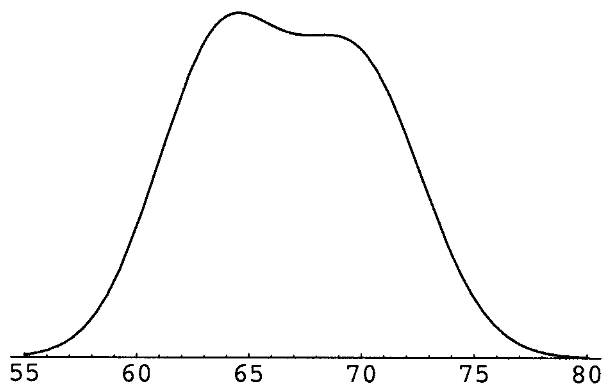

Figure 11. Theoretical distribution of U.S. young adult human height, if males round up to the next inch and females round down.

3.4 Bad Bins

Related to the issue of how students round their heights is the broader issue of placing continuous data into bins. When using a histogram to determine whether a set of data has multiple modes, certain bin widths can hide the bimodal structure of the data. For example, Wand (1997) showed how the S-Plus default bandwidth oversmooths a histogram of British incomes. But we have the opposite problem—the smooth, mound-shaped distribution of Figure 1 being placed into one-inch bins and apparently coming out bimodal. The width of the bins cannot explain that.

3.5 Variability Due to Sampling

The earliest living histogram that we can find (Figure 12) shows the heights of students at Connecticut State Agricultural

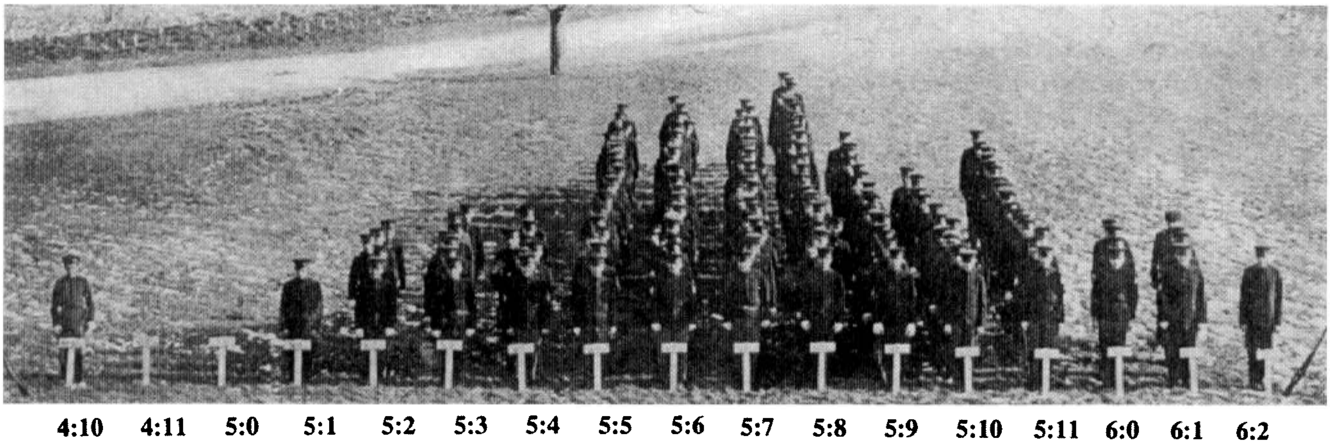

Figure 12. Living histogram of 175 male college students (Blakeslee 1914).

College (now the University of Connecticut). It looks bimodal too, with modal heights at $5' \, 8''$ and $5' \, 11''$—but these students were all men! The explanation is not bad parameters, a bad model, or bad data: "Each of the 175 students shown . . . was measured in his stocking feet and placed in the rank to which his height most nearly corresponded" (Blakeslee 1914). What could explain this?

The distribution of bin counts in a histogram is a multinomial random variable with cell probabilities determined from the underlying distribution. There is more variability in the bins that have the larger counts than in the bins that have smaller counts. A density that is the sum of two normal densities with nearly equal standard deviations and means that are almost two standard deviations apart is mound-shaped with a relatively flat top. In the case of an equal mixture of two normal densities with parameters matching the NHANES data, only six bins ($64''$ to $69''$) hold the middle 53% of the heights and their relative frequencies are quite similar: 0.085, 0.092, 0.093, 0.092, 0.088, and 0.083. With this relatively constant central region, small number of bins, and high variability in their counts, it will not be unusual for samples to give two nonadjacent bins that appear taller than the others.

The destruction of smoothness by sampling is well known. Murphy (1964) constructed 12 histograms for samples of 50 random normal deviates. All but about four have enough irregularity to be reasonably called bimodal. We generated a number of simulated histograms of heights ourselves; very few had a clear unimodal appearance. The literature establishes that using the multimodality of a histogram to infer multimodality of the underlying distribution is unreliable unless the sample size is very large. Both false positive and false negative errors occur frequently.

3.6 Selection Bias

Although the living histogram of heights is often used as an example of a bimodal situation, we have seen only two actual photographs that purport to show this bimodality. One can only wonder how many instructors got their students out on the football field and found themselves with nothing interesting to photograph.

4. SUMMARY

A mixture of two normal densities will not be bimodal unless there is a very large difference between their means, typically larger than the sum of their standard deviations. If male and female heights are even approximately normally distributed with means and standard deviations close to those reported by the NHANES survey, the difference between the means of male and female heights is not enough to produce bimodality. Yet photographs of living histograms do show bimodality.

We have discussed several reasons why a histogram of male and female heights may have a bimodal appearance even though, using reasonable estimates of the parameters, theory says the underlying distribution is mound-shaped. For example, the tendency of students to favor certain values when self-reporting their height may contribute to the phenomenon. And when taking small random samples from a mound-shaped distribution, it is common to get a histogram that looks bimodal.

So what's an instructor to do if he or she decides to jettison this compelling example? We haven't been able to find an equally appealing classroom demonstration of bimodality as the mixture of two populations.

We ran some simulations to see what conditions would be likely to give consistent representations of true bimodality in a classroom experiment. With equal standard deviations and sample sizes of 25 or more per group, a separation in the means of about three standard deviations usually produces good results when using around 15–20 intervals. Occasionally the histogram is too irregular to be clearly identified as bimodal, and in rare instances looks more unimodal than bimodal. Such cases are of course more rare for larger sample sizes than for small ones. Perhaps a reader will be able to provide an easy-to-implement classroom demonstration having approximately this degree of separation between the components.

[Received May 2001. Revised August 2001.]

REFERENCES

Behboodian, J. (1970), "On the Modes of a Mixture of Two Normal Distributions," *Technometrics*, 12, 131–139.

Blakeslee, A. F. (1914), "Corn and Men," *The Journal of Heredity*, 5, 511–518.

Cohen, A. C. (1956), "Compound Normal Distribution" (Advanced Problems and Solutions), *American Mathematical Monthly*, 63, 129.

Crow, J. F. (1997), "Birth Defects, Jimson Weeds and Bell Curves," *Genetics*, 147, 1–6.

Devore, J., and Peck, R. (1997), *Statistics: The Exploration and Analysis of Data*, Belmont, CA: Duxbury.

Eisenberger, I. (1964), "Genesis of Bimodal Distributions," *Technometrics*, 6, 357–364.

Everitt, B. S., and Hand, D. J. (1981), *Finite Mixture Distributions*, London: Chapman and Hall.

The Hartford Courant (1996), "Reaching New Heights," November 23, 1996; photo by K. Hanley.

Helguerro, F. (1904), "Sui Massimi Delle Curve Dimorfiche," *Biometrika*, 3, 85–98.

Hodges, J. L. Jr., Krech, D., and Crutchfield, R. S. (1975), *StatLab: An Empirical Introduction to Statistics*, New York: McGraw-Hill.

Holgersson, M., and Jorner, U. (1978), "Decomposition of a Mixture into Two Normal Components: A Review," *International Journal of Bio-Medical Computing*, 9, 367–392.

Iversen, G. R., and Gergen, M. (1997), *Statistics, the Conceptual Approach*, New York: Springer.

Joiner, B. L. (1975), "Living Histograms," *International Statistical Review*, 3, 339–340.

Kakiuchi, I. (1981), "Unimodality Conditions of the Distribution of a Mixture of Two Distributions," *Kobe University. Mathematics Seminar Notes* 9, 315–325.

Kemperman, J. H. B. (1991), "Mixtures with a Limited Number of Modal Intervals," *The Annals of Statistics*, 19, 2120–2144.

Murphy, E. A. (1964), "One Cause? Many Causes? The Argument from the Bimodal Distribution," *Journal of Chronic Diseases*, 17, 301–324.

Pearson, K. (1894), "Contribution to the Mathematical Theory of Evolution," *Philosophical Transactions of the Royal Statistical Society of London*, Ser. A, 185, 71–110.

Robertson, C. A., and Fryer, J.G. (1969), "Some Descriptive Properties of Normal Mixtures," *Skandinavian Aktuarietidskrift*, 52, 137–146.

Saville, D. J., and Wood, G. R. (1996), *Statistical Methods: A Geometric Primer*, New York: Springer, 66.

Titterington, D. M., Smith, A. F. M., and Makov, U. E. (1985), *Statistical Analysis of Finite Mixture Distributions*, Chichester: Wiley.

U.S. Census Bureau (1999), *Statistical Abstract of the United States: 1999.* Table #243, p. 155. Source of table: U.S. National Center for Health Statistics, unpublished data. Data apparently reorganized from: U.S. Department of Health and Human Services (DHHS), National Center for Health Statistics, Third National Health and Nutrition Examination Survey, 1988-1994, NHANES III Laboratory Data File. Public Use Data File Documentation Number 76200. Hyattsville, MD: Centers for Disease Control and Prevention, 1996.

Wand. M. P. (1997), "Data-Based Choice of Histogram Bin Width," *The American Statistician*, 51, 59–64.

Wild, C. J., and Seber, G. A. F. (2000), *Chance Encounters: A First Course in Data Analysis and Inference*, New York: Wiley, pp. 58-60.