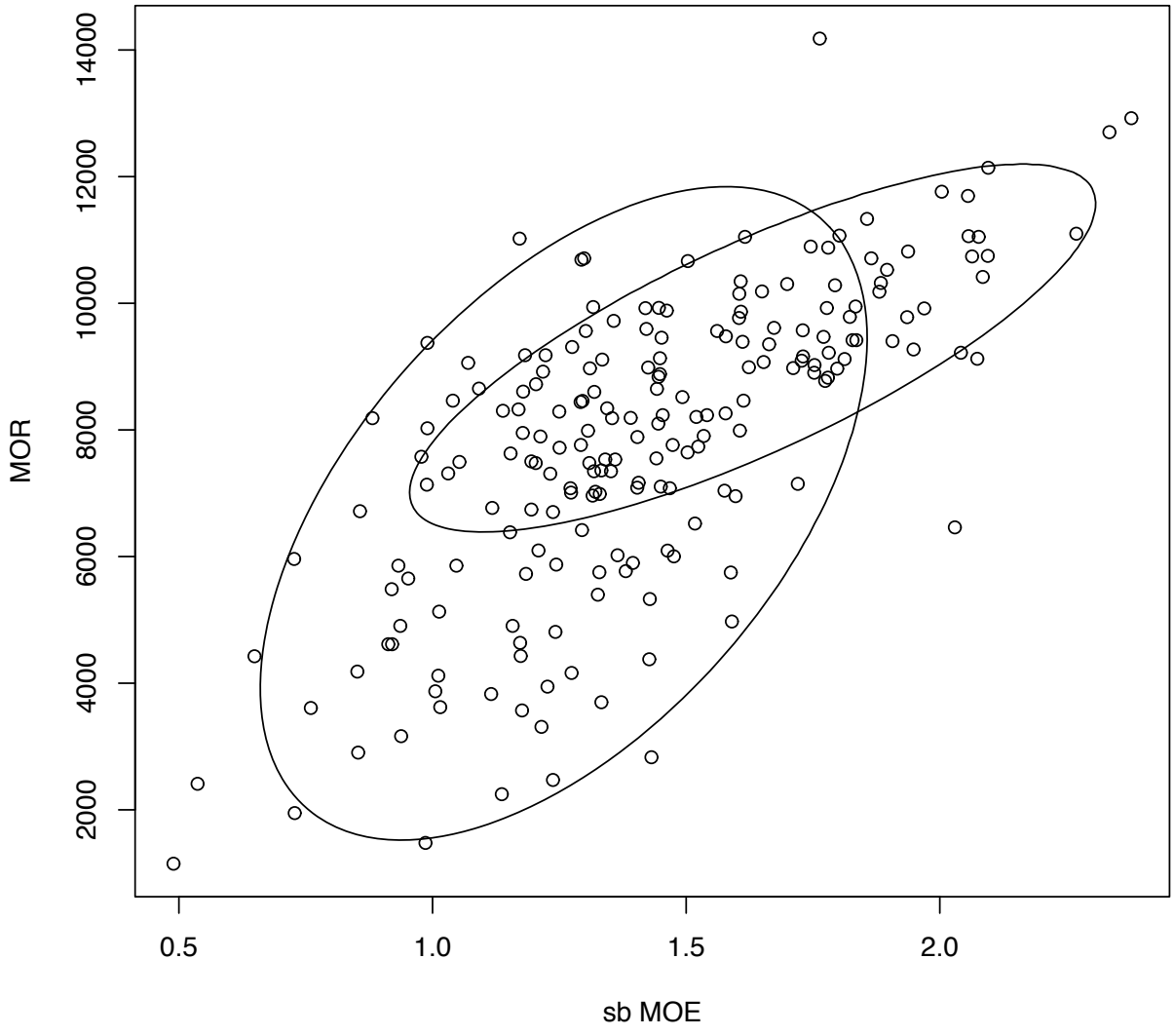
USDA United States Department of Agriculture

A Fit of a Mixture of Bivariate Normals to Lumber Stiffness–Strength Data

Steve P. Verrill
Frank C. Owens
David E. Kretschmann
Rubin Shmulsky



Forest Service | Forest Products Laboratory | Research Paper FPL–RP–696 | August 2018

Abstract

It has been common practice to assume that a two-parameter Weibull probability distribution is suitable for modeling lumber strength properties. In a series of papers published from 2012 to 2018, Verrill et al. demonstrated theoretically and empirically that the modulus of rupture (MOR) distribution of a visual grade of lumber or of lumber that has been "binned" by modulus of elasticity (MOE) is not a two-parameter Weibull. Instead, the tails of the MOR distribution are thinned via "pseudo-truncation." The theoretical portion of Verrill et al.'s argument was based on the assumption of a bivariate normal–Weibull MOE–MOR distribution for the full ("mill run") population of lumber. Verrill et al. felt that it was important to investigate this assumption. In a recent pair of papers, they reported results obtained from a sample of size 200 drawn from a mill run population. They found that normal, lognormal, three-parameter beta, and Weibull distributions did not fit the sample MOR distribution of these data. Instead, it appeared that the MOR data might be fit by a skew normal distribution or a mixture of two univariate normals. In this paper, we investigate whether the joint MOE–MOR data from Verrill et al.'s recent mill run study can be well modeled as a mixture of two bivariate normals.

August 2018

Verrill, Steve P.; Owens, Frank C.; Kretschmann, David E.; Shmulsky, Rubin. 2018. A fit of a mixture of bivariate normals to lumber stiffness–strength data. Research Paper FPL-RP-696. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 44 p.

A limited number of free copies of this publication are available to the public from the Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53726-2398. This publication is also available online at www.fpl.fs.fed.us. Laboratory publications are sent to hundreds of libraries in the United States and elsewhere.

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin.

Keywords: normal distribution, bivariate normal distribution, mixed normal distribution, mixed bivariate normal distribution, skew normal distribution, two-parameter Weibull distribution, three-parameter Weibull distribution, bivariate Gaussian–Weibull distribution, pseudo-truncated Weibull distribution, pseudo-truncated mixed normal distribution, machine stress-rated lumber, MSR lumber, binned MOE lumber, probability density functions, mill run, thin tail, lumber property distribution, chi-squared goodness-of-fit test for a mixture of bivariate normals.

Contents

A Fit of a Mixture of Bivariate Normals to Lumber Stiffness–Strength Data

Steve P. Verrill, Mathematical Statistician
USDA Forest Service, Forest Products Laboratory, Madison, WI

Frank C. Owens, Assistant Research Professor
Department of Sustainable Bioproducts, Mississippi State University, MS

David E. Kretschmann, Research General Engineer (retired[1])
USDA Forest Service, Forest Products Laboratory, Madison, WI

Rubin Shmulsky, Professor and Department Head
Department of Sustainable Bioproducts, Mississippi State University, MS

1 Introduction

In the wood reliability engineering community it has been common practice to model modulus of elasticity (MOE) as a normal distribution and modulus of rupture (MOR) as a normal, lognormal, or two-parameter Weibull distribution (Green and Evans 1987; Evans et al. 1997; ASTM 2017a, 2017b). However, Verrill et al. (2012, 2015) demonstrated theoretically the (in retrospect, obvious) fact that the MOR population associated with a specific visual grade or specific MOE bounds will not have the same theoretical form as the full ("mill run") MOR population from which the visual grade or the "MOE binned" subpopulation is drawn. In particular, they showed that if the mill run population strength distribution is a two-parameter Weibull, then the strength distribution of a visual grade or a grade based on truncated MOE values cannot be a two-parameter Weibull. Instead, it has a "pseudo-truncated Weibull" distribution. (Verrill et al. (2012, 2015) showed mathematically that a pseudo-truncated Weibull has a probability density function (pdf) that *cannot* be represented as a Weibull probability density function.) Pseudo-truncated Weibull or normal or lognormal distributions will have tails that are "thinned" or "tightened" compared to non-pseudo-truncated Weibull, normal, or lognormal distributions. Verrill et al. (2013, 2014, 2018) noted that such tightened tails can be observed in probability plots of In-Grade data. Verrill et al. (2013, 2014, 2018) further established that if one fits a two-parameter Weibull to pseudo-truncated data, estimates of probabilities of breakage when loads are at "allowable limits" can be seriously in error. (See, for example, ASTM (2016) for a discussion of "allowable limits.")

Verrill et al. (2015) started with a bivariate normal–Weibull MOE–MOR distribution and derived the form of the corresponding pseudo-truncated Weibull model for MOR (for sub-populations formed by specimens subject to specific MOE bounds). However, the actual form of the pseudo-truncated MOR values will depend upon the exact relationship between MOE and MOR. If MOE and MOR do not have the particular bivariate normal–Weibull relationship assumed in Verrill et al. (2015), the pdf of a pseudo-truncated MOR population based on MOE bounds will differ from the pseudo-truncated Weibull form derived by Verrill et al. (2015).

The bivariate normal–Weibull assumption upon which Verrill et al. (2015) based their work is "commonly accepted" (that is, many researchers assume that MOE is distributed as a normal

[1] now President, American Lumber Standard Committee, Inc., Frederick, MD

and MOR is distributed as a Weibull — see, for example, ASTM (2017a, 2017b)). However, these assumptions are based on the work of researchers who were fitting distributions to grades of lumber (as in the In-Grade work) rather than to mill run populations.

This fact, and the need for an accurate estimate of the MOR probability density function in reliability calculations, led us to perform the experiment described in Verrill et al. (2017) and Owens et al. (2018) — an experiment designed to yield estimates of mill run bivariate stiffness–strength distributions (which, in turn, yield estimates of pseudo-truncated MOR distributions).

For this initial study, we decided to restrict ourselves to the population of lumber produced at a single mill during a single day. We realized that populations obtained over multiple days from multiple mills in multiple regions would almost certainly have more complicated structures. However, we felt that it would be useful to first address the fundamental question of whether a "simple" population could be modeled by a "simple" distribution such as a normal or two-parameter Weibull. In Verrill et al. (2017) and Owens et al. (2018), we reported stiffness–strength data and distribution fits for a sample of 200 pieces of mill run 2x4 lumber obtained from a single mill on a single day.

We concluded from those analyses that the mill run MOR data were not well fit by normal, lognormal, three-parameter beta, or two- or three-parameter Weibull distributions, but the data might be well fit by a skew normal distribution or a mixture of two normal distributions. We also concluded that mill run stiffness data might be well fit by a single normal distribution or a mixture of two normal distributions.

In section 4.2 of Verrill et al. (2017) we noted that graphical evidence from the 200-piece sample suggested that bivariate stiffness–strength data might have approximately the distribution of a mixture of two bivariate normal distributions. We further noted that this would explain the good separate fits of mixtures of univariate normals to the stiffness data and to the strength data. Some possible causes of such a mixture population include a mixture of trees from a fast-grown plantation stand and a suppressed stand, trees of two separate species, small-diameter trees and large-diameter trees, and 2x4's from the pith region versus 2x4's from the bark region.

In this paper, we take a more rigorous look at modeling mill run stiffness–strength populations by mixtures of bivariate normals.

In Section 2, we review the graphical evidence for mixtures of bivariate normals as stiffness–strength distributions.

In Section 3, we describe goodness-of-fit tests that we performed to test this hypothesis (and we provide a link to a page on our web site at which we provide the Fortran software that we developed to perform the tests).

In Section 4, we describe the effects of pseudo-truncation applied to a mixture of two bivariate normal distributions, and in Section 5, we discuss a series of recent results, including those presented in the current paper, that conflict (to some extent) with current lumber strength standards. We argue that these conflicts suggest that some current lumber strength standards may eventually need to be revised, and/or new strength prediction techniques may need to be developed. We emphasize that we are *not* currently proposing any changes in standards. Any such proposal would need to be based on much additional work.

2 Graphical evidence for a mixture of bivariate normals

The models discussed in section 4.1 of Verrill et al. (2017) were "univariate" models. That is, they were models for the distributions of single variables (sb MOE[1], ecomp[1], dire[1], or MOR[1]). Our univariate work suggested that sb MOE, ecomp, and dire might be modeled as normals or as mixtures of two univariate normals, and that MOR might be modeled by a skew normal distribution or a mixture of two univariate normals.

Because strength and stiffness are positively correlated, there is extra information to be gained by modeling the joint behavior of sb MOE and MOR (or ecomp and MOR or dire and MOR). In this section, we discuss a bivariate model — a mixture of two bivariate normals — that is suggested by the data. We note that if a bivariate stiffness–strength distribution is a mixture of *bivariate* normals, then the corresponding univariate (or "marginal") stiffness and strength distributions will be mixtures of *univariate* normals. (This is well known to statisticians but we provide an "elementary" proof in Appendix A.) Thus, a stiffness–strength distribution that is well modeled by a mixture of bivariate normals would explain the fact that sb MOE, ecomp, dire, and MOR are individually well modeled by mixtures of univariate normals.

What is the graphical evidence for a mixture of bivariate normals? Statisticians know that if variables X and Y have a bivariate normal distribution, then a plot of Y versus X values will be an approximately elliptical cloud of data points. In Figures 1–3, we plot MOR versus sb MOE, MOR versus ecomp, and MOR versus dire, respectively. All three scatter plots suggest that the bivariate stiffness–strength distributions might be well approximated by mixtures of bivariate normal distributions. To investigate this possibility further, we used maximum likelihood methods to fit mixtures of two bivariate normal distributions to the sb MOE–MOR, ecomp–MOR, and dire–MOR data sets. (The density of a mixture of two bivariate normal distributions is provided in Equation (6) in Appendix A.) The programs that we wrote to perform these fits can be found at `http://www1.fpl.fs.fed.us/mix_bivn_code.html`.

In Table 1 we report the parameter estimates from these fits. The lower left cloud in the appropriate MOE–MOR scatter plot (Figures 1–3) corresponds to the "Left bivariate normal" fit in Table 1. The upper right cloud corresponds to the "Right bivariate normal" fit in Table 1.

The fitted $\hat{\mu}_{\mathrm{MOE},1}$, $\hat{\sigma}_{\mathrm{MOE},1}$, $\hat{\rho}_1$, $\hat{\mu}_{\mathrm{MOR},1}$, $\hat{\sigma}_{\mathrm{MOR},1}$, $\hat{\mu}_{\mathrm{MOE},2}$, $\hat{\sigma}_{\mathrm{MOE},2}$, $\hat{\rho}_2$, $\hat{\mu}_{\mathrm{MOR},2}$, $\hat{\sigma}_{\mathrm{MOR},2}$, and $\hat{p}$ in Table 1 correspond to the parameters $\mu_{\mathrm{X}1}$, $\sigma_{\mathrm{X}1}$, ρ_1, $\mu_{\mathrm{Y}1}$, $\sigma_{\mathrm{Y}1}$, $\mu_{\mathrm{X}2}$, $\sigma_{\mathrm{X}2}$, ρ_2, $\mu_{\mathrm{Y}2}$, $\sigma_{\mathrm{Y}2}$, and p of Equation (6) in Appendix A. In particular, $\hat{p}$ is our estimate of the proportion of specimens that come from the lower left bivariate normal population.

The bivariate fits can be used to calculate "probability contours." For example, for a bivariate normal distribution, a 0.90 probability contour is an ellipse centered at the bivariate mean of the distribution that will (over the long run) contain 90% of (X, Y) pairs randomly drawn from the distribution. In each of Figures 4–6 we have superimposed 0.90 probability ellipses corresponding to the two bivariate normal *components* (the component whose probability density function (pdf) is multiplied by p in Equation (6) and the component whose pdf is multiplied by $1 - p$) of the mixed bivariate normal distributions fitted to the sb MOE–MOR, ecomp–MOR, and dire–MOR data clouds displayed in Figures 1–3. In Figures 7–9 we plot 0.90 probability contours for the *full* mixed bivariate normal distributions (as opposed to the 0.90 content ellipses corresponding to the components). The plots do not *prove* anything. However, the plots do lend support to the intuition

[1]sb MOE (static bending MOE) and MOR were measured on an Instron® universal testing machine (Instron Corporation, Norwood, Massachusetts, USA). Ecomp was obtained by measuring transverse vibration via a Metriguard E-computer Model 340 (Metriguard, Pullman, Washington, USA). Dire was obtained by measuring acoustic velocity via a Fibre-gen Director HM200 (Fibre-gen, Christchurch, New Zealand). Details can be found in section 3 of Verrill et al. (2017).

that mill run lumber stiffness–strength distributions will likely be mixtures of distributions. In our case the stiffness–strength distributions appear to be mixtures of two bivariate normal distributions (at least approximately).

In the next section we supplement this "graphical" evidence with results from formal tests of goodness-of-fit.

3 Chi-squared goodness-of-fit tests

We chose to implement a "chi-squared" goodness-of-fit test for a mixture of bivariate normals. Chi-squared goodness-of-fit tests are described in many statistical textbooks. In chapter 3 of D'Agostino and Stephens (1986), David Moore provides a detailed description of "Tests of Chi-Squared Type." In that chapter, Moore suggests that an appropriate number of "cells" lies between $1.88 \times n^{2/5}$ and $3.76 \times n^{2/5}$ where n is the sample size. For a sample of size 200, this suggests a number of cells between 16 and 32. We chose to work with 20 cells, each of which (under the null hypothesis of a mixture of two bivariate normal distributions) contained about 0.05 of the probability. That is, the expected number of observations in each cell was approximately $200 \times 0.05 = 10$. (In D'Agostino and Stephens (1986), Moore recommends that cells should be of approximately equal probability.)

We followed several steps to perform the test:

1. Obtain a maximum likelihood (ML) fit of a mixture of two bivariate normals to the set of 200 stiffness–strength data pairs.

2. Choose a rectangular region (e.g., 0 to 3 by 0 to 15 for a stiffness–strength region [MOE values divided by 1,000,000, MOR values divided by 1,000]) that contains essentially all the probability (e.g., 0.997 and above).

3. Divide this region into 1,000,000 (1,000 by 1,000) rectangles.

4. Take as the estimate of the probability of one of these small rectangles,

   $$\text{(pdf at center of small rectangle)} \times \text{(area of small rectangle)}$$

   (This amounts to numerical integration.)

5. Use the small rectangles to divide the large rectangle into J columns, each of which contains approximately $1/J$ of the probability.

6. Use the small rectangles to divide each of the J columns into I rows (where $J \times I = 20$), each of which contains approximately 0.05 of the probability. The exact probability associated with a cell will be the sum of all the probabilities associated with the small rectangles contained in the cell. For the ijth cell, $j \in \{1, \ldots, J\}$, $i \in \{1, \ldots, I\}$, denote this probability by p_{ij}.

7. Take the chi-squared statistic to be

   $$\chi^2 = \sum_{j=1}^{J} \sum_{i=1}^{I} (O_{ij} - E_{ij})^2 / E_{ij}$$

   where O_{ij} is the observed number of stiffness–strength pairs in the ijth cell and $E_{ij} = p_{ij} \times 200$.

8. Because we used full data maximum likelihood techniques rather than grouped data maximum likelihood techniques to estimate the parameters of the mixture of bivariate normals, the "chi-squared" statistic is not actually distributed as a chi-squared random variable (see Moore's section 3.2.2.1 in D'Agostino and Stephens (1986)). Thus, we needed to do a "parametric bootstrap" (a kind of simulation) to obtain estimates of the p-values associated with the statistic. For each of the chi-squared tests we performed 1000 simulations. In each of these simulations we generated a sample of size 200 from the fitted (to the original stiffness–strength data) mixture of bivariate normals. For each of these simulated data sets, we calculated a chi-squared statistic by the same method used to calculate the chi-squared statistic for the original data. This resulted in the original chi-squared statistic based on the original data, and 1000 chi-squared statistics based on 1000 samples of size 200 drawn from the mixed bivariate normal distribution that was fit to the original data.

9. From these simulations it is possible to obtain an approximate p-value and an "exact" lower bound on the true p-value. If we order the 1000 simulated chi-squared values from smallest to largest, and the original chi-squared value calculated from the real data lies between the mth and $(m+1)$th of the 1000 ordered chi-squared values, then the approximate p-value is $(1000-m)/1000$. Thus, for example, if the original chi-squared statistic lies below only 100 of the 1000 simulated chi-squared values, we would say that the approximate p-value is 0.10. If the original chi-squared statistic lies below only 50 of the 1000 simulated chi-squared values, we would say that the approximate p-value is 0.05.

   In Appendix C we describe how the "exact" lower bound on the p-value is obtained.

For sb MOE, we performed chi-squared goodness-of-fit tests with

- $J = 5, I = 4$
- $J = 4, I = 5$
- $J = 10, I = 2$
- $J = 2, I = 10$

We did this because we wanted to have some assurance that the results were not heavily dependent on how the 20 cells were formed. The conclusions were not dependent on how the 20 cells were formed, so in the ecomp and dire cases, we performed chi-squared goodness-of-fit tests only with $J = 5, I = 4$.

A listing of one version of the program that we wrote to perform the chi-squared tests and simulations can be found at `http://www1.fpl.fs.fed.us/mixbivn.html` (the program also performed maximum likelihood estimation of mixed bivariate normals and produced various plots).

Figures 10–15 display data clouds overlaid with *overall* mixed bivariate normal 0.90 contours (as opposed to the *component* bivariate normal 0.90 contours displayed in Figures 4–6) and the cells that were used in the chi-squared tests.

Tables 2–7 contain the corresponding observed and expected values for the 20 cells in each test.

Table 8 contains the approximate p-values and the "exact" lower bounds on p-values for the various chi-squared goodness-of-fit tests.

Taken together, the plots and the chi-squared tests suggest that we cannot reject the hypothesis that the stiffness–strength values measured for the 200 pieces of lumber are drawn from a mixture of two bivariate normals. (Having said this, we realize that a mixture of bivariate normals model *cannot* be completely correct because it predicts non-zero probabilities for negative stiffness and strength values.)

4 Consequences of pseudo-truncation and a parent population that is a mixture of bivariate normals

Before discussing the implications of a parent population that is a mixture of two bivariate normals, we first need to present some additional plots. (We used a version of the Fortran program available at `http://www1.fpl.fs.fed.us/mixbivn.html` together with the R statistical package (R Core Team (2013)) to produce these plots.) In Figure 16, we plot in solid black the density of the marginal MOR mixture of normals (calculated from the fit of the mixture of bivariate sb MOE–MOR normals). We also include the $p \times f_{\mathrm{MOR,L}}$ (in dashed red) and $(1-p) \times f_{\mathrm{MOR,R}}$ (in dotted blue) components contributed to this density from the "lower left" and "upper right" fitted subpopulations. (Here, the f values denote the probability density functions of the left and right components of the marginal MOR mixture of normals. See Equation (7) in Appendix A. The fitted value for p was 0.55.) From right to left, the dotted blue vertical lines are at the theoretical 0.05, 0.01, 0.001, and 0.0001 quantiles of the right subpopulation. From right to left, the dashed red lines are at the theoretical 0.05, 0.01, 0.001, and 0.0001 quantiles of the left subpopulation (obviously, the model breaks down for the leftmost tail of the left component as it yields negative strength values for the theoretical 0.001 and 0.0001 quantiles). This plot makes clear the substantial left tail differences between the two subpopulations. This should (perhaps) start us worrying. However, these are mill run results. We would expect that by using visual grades or MOE binning, we would be "accounting for" such differences. That is, the lower quality pieces in the left tail of the left population would be screened out.

To investigate this assumption, we performed sb MOE binning on the sample. That is, we looked at the pseudo-truncated MOR population that is obtained by considering only that subset of the mixture of bivariate normals for which the sb MOE value lies between the 0.4 and 0.8 quantiles of the full marginal sb MOE population.

The sb MOE population used in this 0.4,0.8 truncating is illustrated in Figure 17. There we plot in solid black the density of the marginal sb MOE mixture of normals (calculated from the fit of the mixture of bivariate sb MOE–MOR normals). The vertical black lines are at the 0.4 and 0.8 quantiles of this population — 1.332 and 1.721. We also include the $p \times f_{\mathrm{MOE,L}}$ (in dashed red) and $(1-p) \times f_{\mathrm{MOE,R}}$ (in dotted blue) components contributed to this density from the "lower left" and "upper right" fitted subpopulations. (Here, the f values denote the probability density functions of the left and right components of the marginal sb MOE mixture of normals. See Equation (8) in Appendix A.)

Figure 18 includes a scatter plot of MOR versus sb MOE, 0.90 probability content contours for the two bivariate normal components of the fitted mixed bivariate normal model, and dashed vertical lines at the 0.4 and 0.8 sb MOE quantiles estimated from the marginal sb MOE distribution calculated from the fitted mixed bivariate normal model.

Figure 19 illustrates the effect of pseudo-truncation via sb MOE limits on the left ($p = 1$) and right ($p = 0$) MOR subpopulations. The subpopulation pdfs (without the p and $1-p$ multipliers) prior to pseudo-truncation are plotted as dashed lines. Pseudo-truncation via the 0.4, 0.8 quantiles of the marginal sb MOE distribution (the pdf is derived in Appendix B and displayed in Equation (18)) yields the solid lines.

Note that the pseudo-truncation has shifted the left subpopulation to the right (as we would hope) and has narrowed both the left and right subpopulations. However, even after pseudo-truncation via sb MOE limits, the two subpopulations remain quite different. Their differences will be quantified when we discuss Figure 21 below.

The density of the resulting 0.4,0.8 pseudo-truncated (on sb MOE) full MOR population is

plotted in Figure 20 together with dashed vertical lines at its 5th percentile and "allowable property" values. (The fifth is 5.03 and the "allowable property" — 5th/2.1 — is 2.40.) Comparing Figures 16 and 20 (compare, for example, the pdf values at MOR = 5 for the two plots, or the pdf values at MOR = 12 for the two plots), we see that (as we would expect) the pseudo-truncation thins both the left and right tails of the MOR distribution.

(The derivations of the marginal non-pseudo-truncated and pseudo-truncated MOR probability density functions are provided in Appendices A and B.)

In Figure 21, we present plots of the pseudo-truncated MOR pdfs based on $p = 1$ (dashed red) and $p = 0$ (dotted blue) values — that is, pseudo-truncated MOR pdfs derived from the "lower left" and "upper right" bivariate normals in the scatter plots. (These pdfs match the solid red and blue pdfs in Figure 19.) From right to left, the dotted blue vertical lines are at the 0.05, 0.01, 0.001, and 0.0001 quantiles of the right subpopulation. From right to left, the dashed red lines are at the 0.05, 0.01, 0.001, and 0.0001 quantiles of the left subpopulation.

As we noted in our discussion of Figure 19, the means of these two subpopulations are closer to each other than the means of the corresponding non-pseudo-truncated subpopulations plotted in Figure 16. However, it is clear from Figure 21 that even after pseudo-truncation via MOE limits, the remaining left and right subpopulations differ considerably in their strength properties. The means, standard deviations, and 0.0001, 0.001, 0.01, 0.05, and 0.50 quantiles for the subpopulations (left and right non-pseudo-truncated and left and right pseudo-truncated) are presented in Table 9.

5 Discussion

This paper and past work have raised concerns about some aspects of current ASTM lumber strength standards:

1. Visual grade or MOE bin based strength populations from a mill will be pseudo-truncated versions of mill run strength populations. Thus, even if, for example, the mill run MOR distribution of 2x4's produced by a mill were a normal or a lognormal or a Weibull, No. 2 lumber from that mill would *not* be normal, lognormal, or Weibull. It would be "pseudo-truncated" normal, lognormal, or Weibull. This fact and its implications are discussed in detail in Verrill et al. (2012, 2013, 2014, 2015, 2018).

2. The data discussed in the current paper and in Verrill et al. (2017) and Owens et al. (2018) suggest that the mill run MOR distribution at one mill is *not* normal, lognormal, or Weibull. The mill run distribution might be a skew normal or a mixture of two normals. We will be replicating this study as well as investigating mill run strength distributions at three other mills.

3. As suggested in Verrill et al. (2017) and established more rigorously in the current paper, the mill run bivariate sb MOE–MOR distribution (at the mill under investigation) can be well modeled as a mixture of two bivariate normals (with mixture proportions 0.55 and 0.45). These data can be assumed to be derived from a random sample of the 2x4 lumber produced by the mill over a period of several hours. Some possible causes of such a mixture population include a mixture of trees from a fast-grown plantation stand and a suppressed stand, trees of two separate species, small-diameter trees and large-diameter trees, and 2x4's from the pith region versus 2x4's from the bark region. In any event, we have found that the strength properties of the two populations that make up the mixture of two bivariate normals are considerably different. This suggests that unless mixtures and mixture proportions are

constant across mills and over time (this seems unlikely), the quality of output could vary significantly over time and among mills depending upon the particular mixture and mixture proportions. (The variability will surprise no one. The ability to quantify one source of variability — varying mixture proportions — and to potentially model its conseqences is new.)

4. It could be argued that visual grading/MOE binning protects us from the effects of input streams of varying quality. However, after performing virtual MOE binning based on our sample of 200 boards, we saw that we could still have large swings in strength properties depending upon which of the two bivariate normal distributions was providing the boards. We chose as our MOE cutoffs the 40th and 80th percentiles of the combined MOE population. We then calculated the resulting pseudo-truncated normal MOR distributions associated with drawing solely from the "lower quality" (0.55 proportion in our sample) population or solely from the "higher quality" (0.45 proportion in our sample) population. (In the absence of detailed knowledge about the sources of the two populations, it certainly seems conceivable that there might be days [or periods within a day] during which most of the logs processed come from the lower quality population.) We found that *even if we impose the 0.4,0.8 MOE full population cutoffs*, the MOR distribution that will result from drawing solely from the "lower quality" bivariate normal will be *significantly* left-shifted (a significantly lower mean and quantiles) from the MOR distribution that will result from drawing solely from the "higher quality" bivariate normal. This might make one worry about the consistency of a mill's product. (We note that this concern cannot directly apply to MSR lumber. In addition to being based on MOE limits, MSR lumber also involves visual grade restrictions, knot size limitations, and quality assurance via mandated and frequent MOR testing.)

5. The fact that MOE bin and visual-grade based MOR strength distributions are not two-parameter Weibulls (and that assuming that they are can lead to large over- or underestimates of probabilities of failure at allowable property loads — see Verrill et al. 2013, 2014, 2018) suggests that the current ASTM (2017b) LRFD approach (which is based on a two-parameter Weibull strength distribution assumption) might need to be re-evaluated and improved.

We believe that the issues raised in this section may eventually need to be addressed by standards-writing bodies. Currently, we are engaged in additional studies that should give us a better idea of the manner in which bivariate MOE–MOR distributions change from mill to mill and from time to time. These studies should help us determine whether we should be concerned about the validity of some of the reliability models that appear in current lumber standards.

6 Summary

Verrill et al. (2012, 2015) demonstrated theoretically that the MOR population associated with a specific visual grade or specific MOE bounds will not have the same theoretical form as the mill run MOR population from which the visual grade or the MOE binned subpopulation is drawn. In particular, they showed that if the mill run population strength distribution is a two-parameter Weibull, the strength distribution of a visual grade or a grade based on truncated MOE values cannot be a two-parameter Weibull. Instead, it has a "pseudo-truncated Weibull" distribution. Pseudo-truncated Weibull or normal or lognormal distributions will have tails that are "thinned" or "tightened" compared to non-pseudo-truncated Weibull, normal, or lognormal distributions.

Verrill et al. (2012, 2015) started with a bivariate normal–Weibull MOE–MOR distribution and derived the form of the corresponding pseudo-truncated Weibull model for MOR (for subpopulations formed by specimens subject to specific MOE bounds). However, the actual form of the pseudo-truncated MOR values will depend upon the exact relationship between MOE and MOR. If MOE and MOR do not have the particular bivariate normal–Weibull relationship assumed in Verrill et al. (2012, 2015), the pdf of a pseudo-truncated MOR population based on MOE bounds will differ from the pseudo-truncated Weibull form derived by Verrill et al. (2012, 2015).

The bivariate normal–Weibull assumption upon which Verrill et al. based their theoretical work is "commonly accepted" (that is, many researchers assume that MOE is distributed as a normal and MOR is distributed as a Weibull). However, these assumptions are based on the work of researchers who were fitting distributions to grades of lumber (as in the In-Grade work) rather than to mill run populations.

This fact, and the need for an accurate estimate of the MOR probability density function in reliability calculations led us to perform the experiment described in Verrill et al. (2017) and Owens et al. (2018) — an experiment designed to yield estimates of mill run bivariate MOE–MOR distributions (which, in turn, yield estimates of marginal pseudo-truncated MOR distributions).

For this initial study, we restricted ourselves to the population of lumber produced at a single mill during a single day. We concluded from our analyses that the mill run MOR data set was not well fit by normal, lognormal, three-parameter beta, or two- or three-parameter Weibull distributions, but it might be well fit by a skew normal distribution or a mixture of two normal distributions. We also concluded that mill run MOE data might be well fit by a single normal distribution or a mixture of two normal distributions. We reported these results in Verrill et al. (2017) and Owens et al. (2018).

In this paper we have extended the formal *univariate* analyses reported in Verrill et al. (2017) and Owens et al. (2018) to formal analyses of the *bivariate* relationships between measures of MOE and MOR. We conclude that, *for the sample studied*, these relationships are well modeled by mixtures of two bivariate normals. (Some possible sources of two component mixture relationships include a mixture of trees from a fast-grown plantation stand and a suppressed stand, trees of two separate species, small-diameter trees and large-diameter trees, and 2x4's from the pith region versus 2x4's from the bark region.)

These analyses permit us to obtain quantitative estimates of the effects of changes in the mixture of raw materials entering a saw mill. This permits us to conclude that, potentially, variable sources of supply may occasionally lead to reduced reliability results even if we require that specimens meet certain stiffness standards. (Of course, reliability results will improve in the presence of added visual requirements and quality control standards.)

This is clearly very preliminary work that is based on a single sample of size 200 from a single mill on a single day. We are currently engaged in a broader study involving a total of seven additional samples of size 200 from four mills. This study will yield information on strength predictors in addition to stiffness measures. It will also permit us to determine whether skew normal and mixed normal distributions continue to fit relatively well to mill run MOR data, and whether Weibull, normal, and lognormal distributions (for example) continue to fit poorly.

7 References

ASTM (2016). Standard Practice for Establishing Allowable Properties for Visually-Graded Dimension Lumber from In-Grade Tests of Full-Size Specimens, D1990 - 16, *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA.

ASTM (2017a). Standard Practice for Sampling and Data-Analysis for Structural Wood and Wood-Based Products, D2915 - 17, *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA.

ASTM (2017b). Standard Specification for Computing Reference Resistance of Wood-Based Materials and Structural Connections for Load and Resistance Factor Design, D5457 - 17, *Annual Book of ASTM Standards*, ASTM International, West Conshohocken, PA.

D'Agostino, R.B. and Stephens, M.A. (1986). *Goodness-of-fit Techniques.* Marcel Dekker, New York, New York.

Evans, J.W., Johnson, R.A., and Green, D.W. (1997). "Goodness-of-fit tests for two-parameter and three-parameter Weibull distributions," Chapter 9 of *Advances in the theory and practice of statistics: A volume in honor of Samuel Kotz*, New York, NY: John Wiley & Sons, Inc., pages 159-178.

Green, D.W. and Evans, J.W. (1987). *Mechanical properties of visually graded lumber: Volumes 1-8*, Department of Commerce, National Technical Information Service, Springfield, VA, 3515 pages.

Owens, F.C., Verrill, S.P., Shmulsky, R., and Kretschmann, D.E. (2018), "Distributions of MOE and MOR in a Full Lumber Population," *Wood and Fiber Science*, **50**(3), pp. 265-269.

R Core Team (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`

Verrill, S., Evans, J.W., Kretschmann, D.E., and Hatfield, C.A. (2012). "Asymptotically Efficient Estimation of a Bivariate Gaussian–Weibull Distribution and an Introduction to the Pseudo-truncated Weibull." Research Paper FPL-RP-666. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 76 pages.

Verrill, S., Evans, J.W., Kretschmann, D.E., and Hatfield, C.A. (2013). "An Evaluation of a Proposed Revision of the ASTM D 1990 Grouping Procedure." Research Paper FPL-RP-671. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 34 pages.

Verrill, S., Evans, J.W., Kretschmann, D.E., and Hatfield, C.A. (2014), "Reliability Implications in Wood Systems of a Bivariate Gaussian–Weibull Distribution and the Associated Univariate Pseudo-truncated Weibull," *ASTM Journal of Testing and Evaluation*, **42**, Number 2, pages 412-419.

Verrill, S., Evans, J.W., Kretschmann, D.E., and Hatfield, C.A. (2015), "Asymptotically Efficient Estimation of a Bivariate Gaussian–Weibull Distribution and an Introduction to the Pseudo-truncated Weibull," *Communications in Statistics – Theory and Methods*, **44**, pp. 2957-2975.

Verrill, S.P., Owens, F.C., Kretschmann, D.E., and Shmulsky, R. (2017), "Statistical Models for the Distributions of Modulus of Elasticity and Modulus of Rupture in Lumber with Implications for Reliability Calculations." Research Paper FPL-RP-692. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 51 pages.

Verrill, S.P., Owens, F.C., Kretschmann, D.E., Shmulsky, R., and Brown, L.S. (2018), "Visual and MSR Grades of Lumber Are Not 2-parameter Weibulls and Why It Matters," draft, `http://www1.fpl.fs.fed.us/weib2.pdf`.

8 Appendix A — Proof that the marginal distributions of a bivariate normal are univariate normals, and the marginal distributions of a mixture of two bivariate normals are mixtures of two univariate normals

Statisticians know that the marginal distributions of a bivariate normal distribution are univariate normals. (See, for example, Roussas (1973).) For the reader's convenience, we supply the standard "elementary" proof (it relies only upon calculus) below.

The probability distribution function of a bivariate normal distribution is

$$f(x, y; \mu_X, \sigma_X, \rho, \mu_Y, \sigma_Y) = \frac{1}{2\pi} \times \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}} \times \exp(-\arg) \tag{1}$$

where

$$\arg = \left[\left(\frac{x-\mu_X}{\sigma_X} \right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X} \right) \left(\frac{y-\mu_Y}{\sigma_Y} \right) + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 \right] \div (2(1-\rho^2))$$

and μ_X, σ_X are the mean and standard deviation of X; ρ is the correlation between X and Y; and μ_Y, σ_Y are the mean and standard deviation of Y.

Thus

$$\begin{aligned} \text{Prob}(Y \le y) &= \int_{-\infty}^{y} \int_{-\infty}^{\infty} f(s, t; \mu_X, \sigma_X, \rho, \mu_Y, \sigma_Y)\, ds\, dt \\ &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma_X \sqrt{1-\rho^2}} \times \exp(-\arg) \\ &\quad \times \exp\left(\rho^2 \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2)) \right) \\ &\quad \times \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2)) \right) ds\, dt \end{aligned} \tag{2}$$

where

$$\arg = \left(\frac{s-\mu_X}{\sigma_X} - \rho \times \frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2))$$

or

$$
\begin{aligned}
\text{Prob}(Y \le y) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_X \sqrt{1-\rho^2}} \exp\left(- \left(\frac{s-\mu_X}{\sigma_X} - \rho \times \frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2)) \right) ds\, dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_X \sqrt{1-\rho^2}} \exp\left(- \left(\frac{s}{\sigma_X} - \rho \times \frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2)) \right) ds\, dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left(- \left(s - \rho \times \frac{t-\mu_Y}{\sigma_Y} \right)^2 \div (2(1-\rho^2)) \right) ds\, dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho^2}} \exp\left(-s^2/(2(1-\rho^2)) \right) ds\, dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \\
&\quad \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-s^2/2 \right) ds\, dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{t-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) dt \qquad (3)
\end{aligned}
$$

That is, the marginal pdf of Y is

$$\frac{d}{dy} \text{Prob}(Y \le y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_Y} \exp\left(- \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 \div 2 \right) \qquad (4)$$

Similarly, the marginal pdf of X is

$$\frac{d}{dx} \text{Prob}(X \le x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_X} \exp\left(- \left(\frac{x-\mu_X}{\sigma_X} \right)^2 \div 2 \right) \qquad (5)$$

More generally, by essentially the same argument, we can show that if we begin with the mixed bivariate normal pdf

$$
\begin{aligned}
&f_{\text{M}}(x, y; \mu_{X1}, \sigma_{X1}, \rho_1, \mu_{Y1}, \sigma_{Y1}, p, \mu_{X2}, \sigma_{X2}, \rho_2, \mu_{Y2}, \sigma_{Y2}) \qquad (6) \\
&= p \times f(x, y; \mu_{X1}, \sigma_{X1}, \rho_1, \mu_{Y1}, \sigma_{Y1}) + (1-p) \times f(x, y; \mu_{X2}, \sigma_{X2}, \rho_2, \mu_{Y2}, \sigma_{Y2})
\end{aligned}
$$

where the $f(x, y)$'s on the right side of the equation are given by Equation (1), we end up with

$$p \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y1}} \exp\left(- \left(\frac{y-\mu_{Y1}}{\sigma_{Y1}} \right)^2 \div 2 \right) + (1-p) \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y2}} \exp\left(- \left(\frac{y-\mu_{Y2}}{\sigma_{Y2}} \right)^2 \div 2 \right) \qquad (7)$$

as the marginal Y pdf, and

$$p \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{X1}} \exp\left(- \left(\frac{x - \mu_{X1}}{\sigma_{X1}} \right)^2 \div 2 \right) + (1-p) \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{X2}} \exp\left(- \left(\frac{x - \mu_{X2}}{\sigma_{X2}} \right)^2 \div 2 \right) \quad (8)$$

as the marginal X pdf.

9 Appendix B — The cumulative distribution function (cdf) and probability density function (pdf) of a pseudo-truncated mixture of bivariate normals

Let (X, Y) have a mixture of bivariate normals distribution. That is, their bivariate pdf function has the form

$$\begin{aligned} & f_{\mathrm{M}}(x, y; \mu_{X1}, \sigma_{X1}, \rho_1, \mu_{Y1}, \sigma_{Y1}, p, \mu_{X2}, \sigma_{X2}, \rho_2, \mu_{Y2}, \sigma_{Y2}) \\ & = p \times f(x, y; \mu_{X1}, \sigma_{X1}, \rho_1, \mu_{Y1}, \sigma_{Y1}) + (1-p) \times f(x, y; \mu_{X2}, \sigma_{X2}, \rho_2, \mu_{Y2}, \sigma_{Y2}) \end{aligned} \quad (9)$$

where p is the mixing fraction (at any given "draw", we draw from population 1 with probability p and from population 2 with probability $1-p$); $\mu_{X1}, \sigma_{X1}, \rho_1, \mu_{Y1}, \sigma_{Y1}$ are the parameters of the population 1 bivariate normal; $\mu_{X2}, \sigma_{X2}, \rho_2, \mu_{Y2}, \sigma_{Y2}$ are the parameters of the population 2 bivariate normal; and $f(w, z; \mu_W, \sigma_W, \rho, \mu_Z, \sigma_Z)$ is the pdf of a single bivariate normal given by

$$f(w, z; \mu_W, \sigma_W, \rho, \mu_W, \sigma_Z) = \frac{1}{2\pi} \times \frac{1}{\sigma_W \sigma_Z \sqrt{1-\rho^2}} \times \exp(-\arg) \quad (10)$$

where

$$\arg = \left[\left(\frac{w - \mu_W}{\sigma_W} \right)^2 - 2\rho \left(\frac{w - \mu_W}{\sigma_W} \right) \left(\frac{z - \mu_Z}{\sigma_Z} \right) + \left(\frac{z - \mu_Z}{\sigma_Z} \right)^2 \right] \div (2(1-\rho^2))$$

and μ_W, σ_W are the mean and standard deviation of W; ρ is the correlation between W and Z; and μ_Z, σ_Z are the mean and standard deviation of Z.

Then, *by definition*, the marginal cdf of a pseudo-truncated (truncated on X) mixture of bivariate normals is given by

$$\begin{aligned} F_{\mathrm{PT}}(y) &= \mathrm{Prob}(Y \le y | c_{\mathrm{L}} \le X \le c_{\mathrm{U}}) \\ &= \mathrm{Prob}(Y \le y \text{ and } c_{\mathrm{L}} \le X \le c_{\mathrm{U}}) / \mathrm{Prob}(c_{\mathrm{L}} \le X \le c_{\mathrm{U}}) \end{aligned} \quad (11)$$

This is the probability that $Y \le y$ given that X lies between c_{L} and c_{U} (for example, the probability that the MOR lies below y given that the MOE lies between the 40th and 80th percentiles of the MOE distribution).

F_{PT} will have 13 parameters — those of X, Y and c_{L} and c_{U}. These latter two parameters will be *known* in the case of MOE binning.

In Appendix A we (essentially) showed that the marginal distributions of a non-pseudo-truncated mixture of bivariate normals will be mixtures of univariate normals.

In the material below we obtain the marginal distribution of the pseudo-truncated Y portion of a mixture of bivariate normals after a truncation of the correlated X portion.

9.1 Denominator of Equation (11)

In the pseudo-truncated case, we have

$$\begin{aligned}
\text{Prob}(c_{\text{L}} \le X \le c_{\text{U}}) &= \text{Prob}\,(X \text{ drawn from pop 1 and } c_{\text{L}} \le X_1 \le c_{\text{U}}) \\
&\quad + \text{Prob}\,(X \text{ drawn from pop 2 and } c_{\text{L}} \le X_2 \le c_{\text{U}}) \\
&= p \times \text{Prob}\left(\frac{c_{\text{L}} - \mu_{X1}}{\sigma_{X1}} \le \frac{X_1 - \mu_{X1}}{\sigma_{X1}} \le \frac{c_{\text{U}} - \mu_{X1}}{\sigma_{X1}}\right) \\
&\quad + (1-p) \times \text{Prob}\left(\frac{c_{\text{L}} - \mu_{X2}}{\sigma_{X2}} \le \frac{X_2 - \mu_{X2}}{\sigma_{X2}} \le \frac{c_{\text{U}} - \mu_{X2}}{\sigma_{X2}}\right) \\
&= p \times \left(\Phi\left(\frac{c_{\text{U}} - \mu_{X1}}{\sigma_{X1}}\right) - \Phi\left(\frac{c_{\text{L}} - \mu_{X1}}{\sigma_{X1}}\right)\right) \\
&\quad + (1-p) \times \left(\Phi\left(\frac{c_{\text{U}} - \mu_{X2}}{\sigma_{X2}}\right) - \Phi\left(\frac{c_{\text{L}} - \mu_{X2}}{\sigma_{X2}}\right)\right) \qquad (12)
\end{aligned}$$

where Φ denotes the N(0,1) cumulative distribution function.

9.2 Numerator of Equation (11)

We have

$$\begin{aligned}
\text{Prob}(Y \le y \text{ and } c_{\text{L}} \le X \le c_{\text{U}}) &= p \times \text{Prob}(Y_1 \le y \text{ and } c_{\text{L}} \le X_1 \le c_{\text{U}}) \\
&\quad + (1-p) \times \text{Prob}(Y_2 \le y \text{ and } c_{\text{L}} \le X_2 \le c_{\text{U}}) \qquad (13)
\end{aligned}$$

and for bivariate normal X, Y, we have (following the same kind of argument as that made in Appendix A)

$$\begin{aligned}
\text{Prob}(Y_1 \le y \text{ and } c_{\text{L}} \le X_1 \le c_{\text{U}}) &= \int_{-\infty}^{y} \int_{c_{\text{L}}}^{c_{\text{U}}} f(s,t;\mu_{X1},\sigma_{X1},\rho_1,\mu_{Y1},\sigma_{Y1})\,ds\,dt \\
&= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y1}} \int_{c_{\text{L}}}^{c_{\text{U}}} \frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma_{X1}\sqrt{1-\rho_1^2}} \times \exp(-\text{arg}) \\
&\quad \times \exp\left(\rho_1^2 \left(\frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)^2 \div (2(1-\rho_2^2))\right) \\
&\quad \times \exp\left(-\left(\frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)^2 \div (2(1-\rho_1^2))\right) ds\,dt
\end{aligned}$$

where

$$\text{arg} = \left(\frac{s-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)^2 \div (2(1-\rho_1^2))$$

or

$$\begin{aligned}
\text{Prob}(Y_1 \le y \text{ and } c_{\text{L}} \le X_1 \le c_{\text{U}}) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y1}} \exp\left(-\left(\frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)^2 \div 2\right) \\
&\quad \times R_1(t)\,dt \qquad (14)
\end{aligned}$$

where

$$
\begin{aligned}
R_1(t) &= \int_{c_{\mathrm{L}}}^{c_{\mathrm{U}}} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{X1}\sqrt{1-\rho_1^2}} \exp\left(-\left(\frac{s-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}} \right)^2 \div (2(1-\rho_1^2)) \right) ds \\
&= \int_{c_{\mathrm{L}}-\mu_{X1}}^{c_{\mathrm{U}}-\mu_{X1}} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{X1}\sqrt{1-\rho_1^2}} \exp\left(-\left(\frac{s}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}} \right)^2 \div (2(1-\rho_1^2)) \right) ds \\
&= \int_{\frac{c_{\mathrm{L}}-\mu_{X1}}{\sigma_{X1}}}^{\frac{c_{\mathrm{U}}-\mu_{X1}}{\sigma_{X1}}} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho_1^2}} \exp\left(-\left(s - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}} \right)^2 \div (2(1-\rho_1^2)) \right) ds \\
&= \int_{\frac{c_{\mathrm{L}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}}}^{\frac{c_{\mathrm{U}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}}} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1-\rho_1^2}} \exp\left(-s^2/(2(1-\rho_1^2)) \right) ds \\
&= \int_{\left(\frac{c_{\mathrm{L}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)/\sqrt{1-\rho_1^2}}^{\left(\frac{c_{\mathrm{U}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}}\right)/\sqrt{1-\rho_1^2}} \frac{1}{\sqrt{2\pi}} \exp\left(-s^2/2 \right) ds \qquad (15) \\
&= \Phi\left(\left(\frac{c_{\mathrm{U}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}} \right) \div \sqrt{1-\rho_1^2} \right) \\
&\quad -\Phi\left(\left(\frac{c_{\mathrm{L}}-\mu_{X1}}{\sigma_{X1}} - \rho_1 \times \frac{t-\mu_{Y1}}{\sigma_{Y1}} \right) \div \sqrt{1-\rho_1^2} \right)
\end{aligned}
$$

where Φ denotes the N(0,1) cumulative distribution function.

Similarly,

$$
\begin{aligned}
\mathrm{Prob}(Y_2 \le y \text{ and } c_{\mathrm{L}} \le X_2 \le c_{\mathrm{U}}) &= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y2}} \exp\left(-\left(\frac{t-\mu_{Y2}}{\sigma_{Y2}} \right)^2 \div 2 \right) \\
&\quad \times R_2(t)\, dt \qquad (16)
\end{aligned}
$$

where

$$
\begin{aligned}
R_2(t) &= \Phi\left(\left(\frac{c_{\mathrm{U}}-\mu_{X2}}{\sigma_{X2}} - \rho_2 \times \frac{t-\mu_{Y2}}{\sigma_{Y2}} \right) \div \sqrt{1-\rho_2^2} \right) \\
&\quad - \Phi\left(\left(\frac{c_{\mathrm{L}}-\mu_{X2}}{\sigma_{X2}} - \rho_2 \times \frac{t-\mu_{Y2}}{\sigma_{Y2}} \right) \div \sqrt{1-\rho_2^2} \right) \qquad (17)
\end{aligned}
$$

From results (11)–(17), we can calculate the marginal distribution function, $F_{\mathrm{PT}}(y)$, of a pseudo-truncated mixture of bivariate normals. By taking the derivative of $F_{\mathrm{PT}}(y)$ with respect to y, we get the pdf at y. That is,

$$
\begin{aligned}
f_{\mathrm{PT}}(y) &= \frac{dF_{\mathrm{PT}}(y)}{dy} \\
&= \left(\frac{d\,\mathrm{Prob}(Y \le y \text{ and } c_{\mathrm{L}} \le X \le c_{\mathrm{U}})}{dy} \right) \div \mathrm{Prob}(c_{\mathrm{L}} \le X \le c_{\mathrm{U}}) \qquad (18)
\end{aligned}
$$

where the numerator of Equation (18) is

$$
p \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{Y1}} \exp\left(-\left(\frac{y-\mu_{Y1}}{\sigma_{Y1}} \right)^2 \div 2 \right) \times R_1(y)
$$

$$+(1-p)\times\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_{Y2}}\exp\left(-\left(\frac{y-\mu_{Y2}}{\sigma_{Y2}}\right)^2\div 2\right)\times R_2(y)$$

where $R_1(y)$ and $R_2(y)$ are given by Equations (15) and (17), and the denominator of Equation (18) is

$$p\times\left(\Phi\left(\frac{c_{\mathrm{U}}-\mu_{X1}}{\sigma_{X1}}\right)-\Phi\left(\frac{c_{\mathrm{L}}-\mu_{X1}}{\sigma_{X1}}\right)\right)+(1-p)\times\left(\Phi\left(\frac{c_{\mathrm{U}}-\mu_{X2}}{\sigma_{X2}}\right)-\Phi\left(\frac{c_{\mathrm{L}}-\mu_{X2}}{\sigma_{X2}}\right)\right)$$

10 Appendix C — "Exact" lower bound on goodness-of-fit p-value

In Section 2 we described a "chi-sqared goodness-of-fit test" for a mixture of bivariate normals. The approximate p-value for the test was taken to be $(1000-m)/1000$ where the test statistic based on the original data lay between the mth and $m+1$th order statistics of the 1000 chi-squared values obtained in the 1000 simulation trials.

We reported these approximate p-values for six tests in Table 8. In Table 8 we also reported "exact" lower bounds on these p-values. In this appendix we describe the manner in which these lower bounds were calculated.

As we noted in Section 3, because we used full data maximum likelihood techniques, the "chi-squared" statistics will not actually have chi-squared distributions. They will have some other distribution F. We approximate this distribution by the empirical distribution obtained from the 1000 trials. If $x_{(1)}\le\ldots\le x_{(1000)}$ are the order statistics obtained in the 1000 trials, then the empirical approximation, F_{1000}, distribution is the step function that jumps by $1/1000$ at each of the 1000 $x_{(i)}$ order statistics.

If we knew the true statistic distribution F and we observed a test statistic value equal to x, then the p-value would be $1-F(x)$.

However, we have only the approximation F_{1000} based on the 1000 trials, and given an observed test statistic equal to x, the approximate p-value provided in Table 8 is $1-F_{1000}(x)$ or $1-m/1000$ for x between the mth and $m+1$th order statistics, $x_{(m)}$ and $x_{(m+1)}$.

Because $x<x_{(m+1)}$, $F(x)\le F(x_{(m+1)})$. If we can get an upper bound on $F(x_{(m+1)})$, we have an upper bound on $F(x)$ and thus a lower bound on the exact p-value $1-F(x)$.

Let ξ_q denote the qth quantile of F. That is, $F(\xi_q)=q$. We have

$$\begin{aligned}\mathrm{Prob}(x_{(m+1)}\le\xi_q) &= \mathrm{Prob}(\text{exactly } m+1 \text{ of the } x_{(i)}\le\xi_q)\\ &\quad+\mathrm{Prob}(\text{exactly } m+2 \text{ of the } x_{(i)}\le\xi_q)\\ &\quad+\ldots\\ &\quad+\mathrm{Prob}(\text{exactly } 1000 \text{ of the } x_{(i)}\le\xi_q)\\ &= \binom{1000}{m+1}q^{m+1}(1-q)^{1000-(m+1)}+\ldots+\binom{1000}{1000}q^{1000}(1-q)^0\end{aligned}\tag{19}$$

This sum can be obtained via the incomplete beta function.
At `http://www1.fpl.fs.fed.us/mixbivn.c.html` we provide a listing of a program that made use of the incomplete beta function to find the smallest q for which this probability is 1. That tells us that, with certainty, $F(x)\le F(x_{(m+1)})\le F(\xi_q)=q$ so

$$\text{p-value}=1-F(x)\ge 1-q$$

This provides the lower bound on the p-value provided in Table 8.

MOE	Left bivariate normal					Right bivariate normal					
	$\hat{\mu}_{\text{MOE},1}$	$\hat{\sigma}_{\text{MOE},1}$	$\hat{\rho}_1$	$\hat{\mu}_{\text{MOR},1}$	$\hat{\sigma}_{\text{MOR},1}$	$\hat{\mu}_{\text{MOE},2}$	$\hat{\sigma}_{\text{MOE},2}$	$\hat{\rho}_2$	$\hat{\mu}_{\text{MOR},2}$	$\hat{\sigma}_{\text{MOR},2}$	$\hat{p}$
sb MOE	1.26	0.28	0.54	6.68	2.40	1.63	0.32	0.79	9.29	1.35	0.55
ecomp	1.46	0.29	0.39	6.49	2.40	1.79	0.37	0.65	9.23	1.36	0.51
dire	1.41	0.28	0.45	6.68	2.45	1.76	0.35	0.63	9.25	1.32	0.55

Table 1: Parameter estimates from mixed bivariate normal fits. "sb MOE" denotes static bending MOE, "ecomp" denotes E-computer MOE, and "dire" denotes Director MOE. The MOE data were in pounds per square inch divided by 1,000,000. The MOR data were in pounds per square inch divided by 1,000.

Column	Row	Observed	Expected
1	1	10	10.038
1	2	8	10.045
1	3	6	10.008
1	4	10	10.475
2	1	11	10.026
2	2	11	10.014
2	3	11	10.114
2	4	18	9.998
3	1	10	10.012
3	2	14	10.105
3	3	9	10.104
3	4	9	10.214
4	1	9	10.067
4	2	6	10.146
4	3	10	10.079
4	4	6	10.000
5	1	15	10.021
5	2	8	10.140
5	3	12	10.128
5	4	7	8.266

Table 2: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. sb MOE–MOR data. 5 columns, 4 rows per column.

Column	Row	Observed	Expected
1	1	12	10.009
1	2	10	10.105
1	3	6	10.009
1	4	5	10.024
1	5	14	10.106
2	1	7	10.013
2	2	14	10.080
2	3	16	10.121
2	4	9	10.079
2	5	10	9.935
3	1	12	10.073
3	2	11	10.180
3	3	7	10.152
3	4	7	10.029
3	5	10	9.598
4	1	11	10.168
4	2	13	10.008
4	3	7	10.009
4	4	12	10.017
4	5	7	9.286

Table 3: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. sb MOE–MOR data. 4 columns, 5 rows per column.

Column	Row	Observed	Expected
1	1	12	10.054
1	2	10	10.250
2	1	6	10.029
2	2	6	10.233
3	1	11	10.012
3	2	14	10.277
4	1	14	10.047
4	2	14	10.603
5	1	10	10.019
5	2	7	10.527
6	1	12	10.055
6	2	11	10.516
7	1	12	10.006
7	2	8	10.556
8	1	6	10.082
8	2	6	10.166
9	1	12	10.084
9	2	11	10.164
10	1	13	10.041
10	2	5	6.276

Table 4: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. sb MOE–MOR data. 10 columns, 2 rows per column.

Column	Row	Observed	Expected
1	1	10	10.076
1	2	13	10.026
1	3	6	10.097
1	4	12	10.010
1	5	4	10.222
1	6	10	10.208
1	7	15	10.210
1	8	12	10.229
1	9	10	10.117
1	10	11	9.287
2	1	12	10.038
2	2	9	10.190
2	3	7	10.047
2	4	11	10.348
2	5	13	10.419
2	6	10	10.151
2	7	10	10.183
2	8	6	10.029
2	9	13	10.012
2	10	6	8.102

Table 5: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. sb MOE–MOR data. 2 columns, 10 rows per column.

Column	Row	Observed	Expected
1	1	9	10.064
1	2	8	10.039
1	3	9	10.069
1	4	12	10.068
2	1	10	10.033
2	2	10	10.093
2	3	9	10.022
2	4	11	10.570
3	1	12	10.002
3	2	15	10.002
3	3	12	10.004
3	4	3	10.337
4	1	13	10.083
4	2	8	10.096
4	3	11	10.055
4	4	14	9.894
5	1	8	10.114
5	2	8	10.121
5	3	10	10.056
5	4	8	8.280

Table 6: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. ecomp–MOR data. 5 columns, 4 rows per column.

Column	Row	Observed	Expected
1	1	8	10.023
1	2	9	10.015
1	3	4	10.069
1	4	15	10.414
2	1	13	10.031
2	2	11	10.042
2	3	11	10.036
2	4	11	10.063
3	1	12	10.074
3	2	14	10.032
3	3	13	10.061
3	4	8	10.221
4	1	10	10.072
4	2	6	10.139
4	3	7	10.120
4	4	10	9.981
5	1	10	10.148
5	2	9	10.115
5	3	9	10.026
5	4	10	8.320

Table 7: "Chi-squared" goodness-of-fit test results. Actual and expected counts for the 20 cells. dire–MOR data. 5 columns, 4 rows per column.

Stiffness	Columns	Rows	GoF Statistic value	Approximate p-value	"Exact" lower bound on p-value
sb MOE	5	4	17.36	0.204	0.15
	4	5	17.15	0.230	0.18
	10	2	15.36	0.433	0.37
	2	10	14.79	0.429	0.36
ecomp	5	4	13.60	0.455	0.39
dire	5	4	11.80	0.634	0.57

Table 8: "Chi-squared goodness-of-fit test" results. We are testing the null hypothesis that the stiffness–strength distribution is a mixture of two bivariate normal distributions.

Type	Subpopulation	Mean	Standard deviation	Quantiles				
				.0001	.001	.01	.05	.50
non-PT	left	6.68	2.40	-2.26	-.75	1.09	2.73	6.68
	right	9.29	1.35	4.26	5.11	6.14	7.07	9.29
0.4,0.8 PT	left	7.71	2.09	-.031	1.27	2.86	4.28	7.71
	right	8.97	0.92	5.59	6.15	6.84	7.46	8.97

Table 9: Means, standard deviations, and quantiles of marginal MOR left and right subpopulations obtained from non-pseudo-truncated and pseudo-truncated mixtures of bivariate normal distributions fitted to the sample of 200 sb MOE–MOR data pairs.

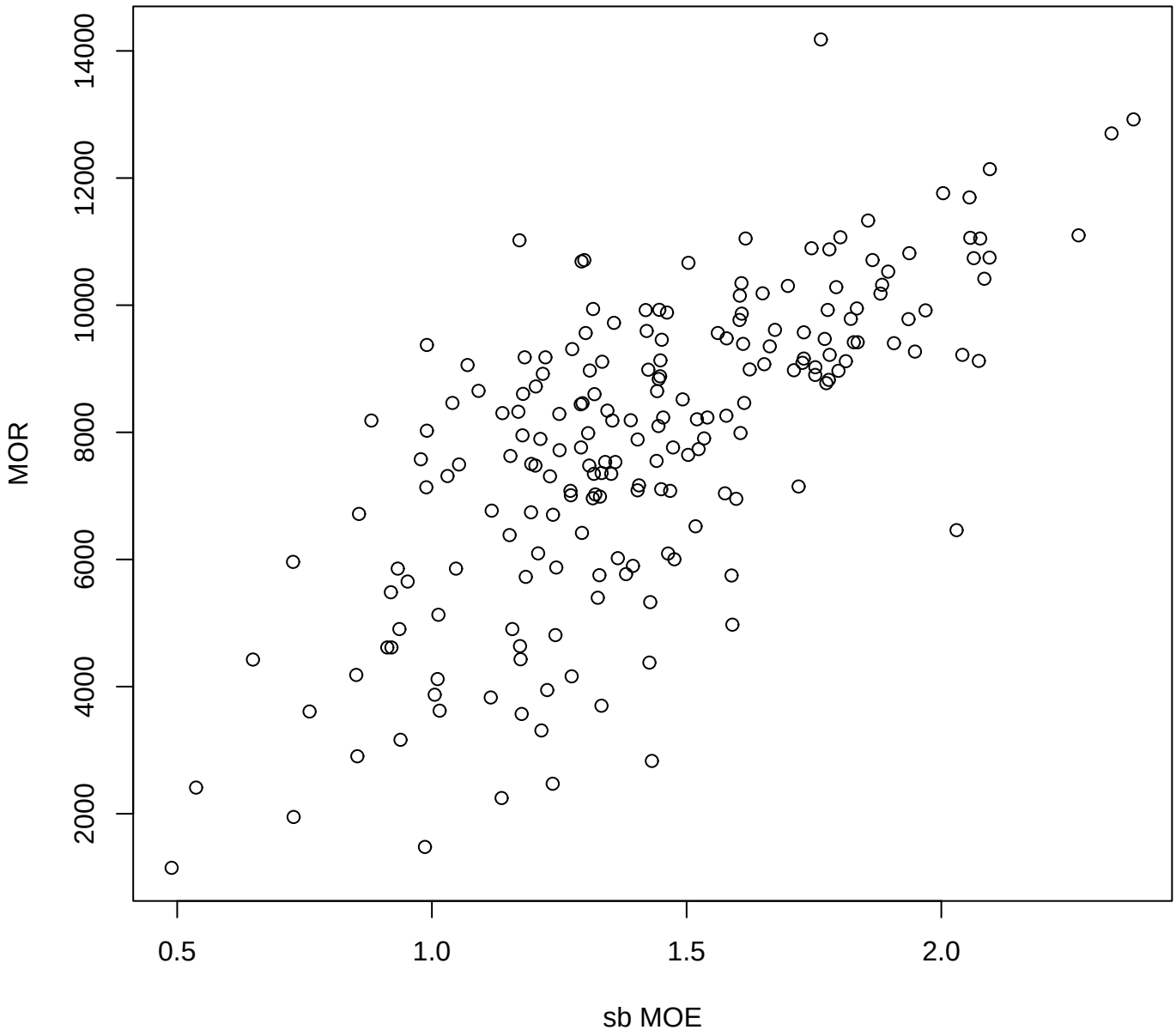


Figure 1: MOR versus sb MOE.

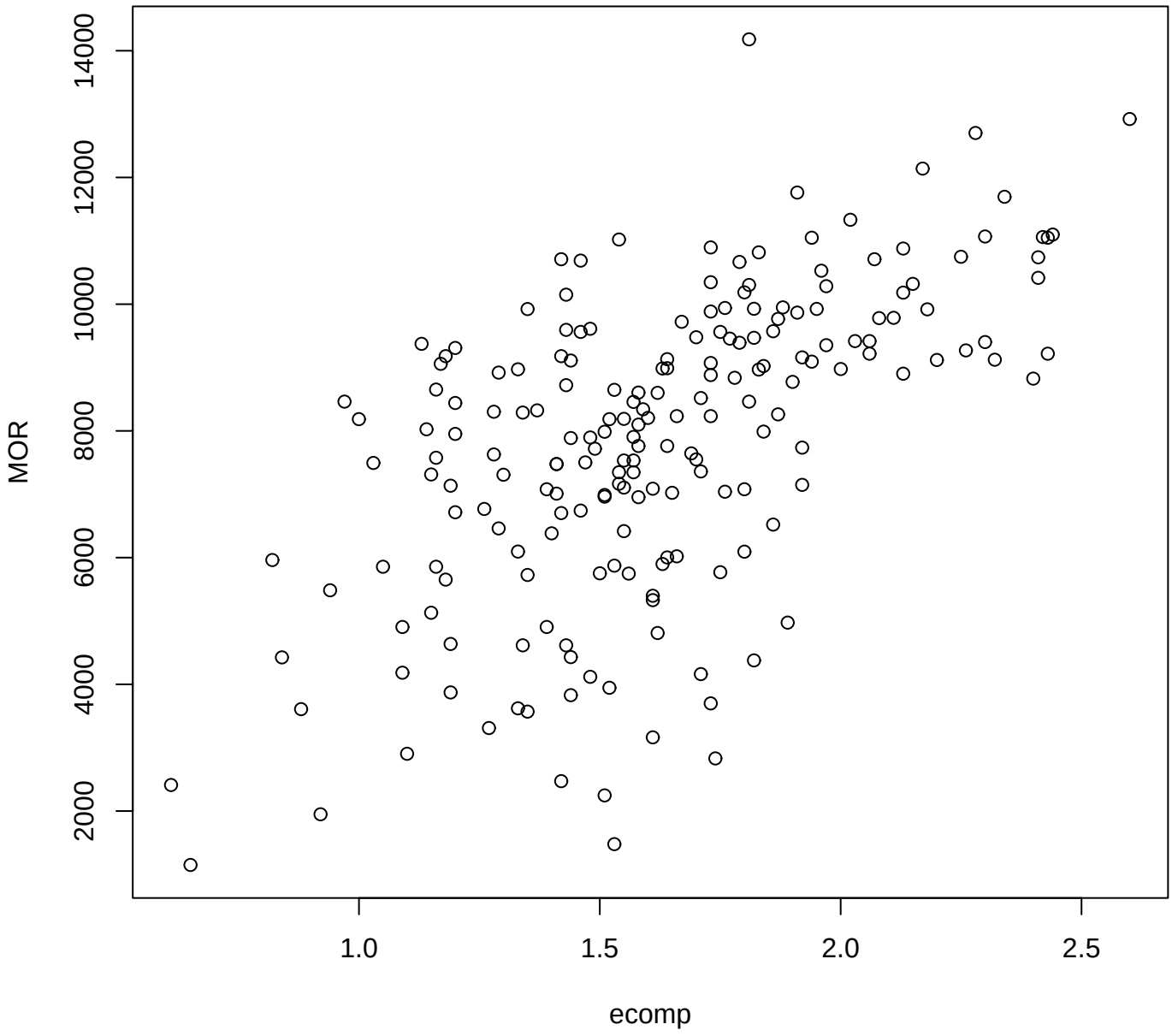


Figure 2: MOR versus ecomp.

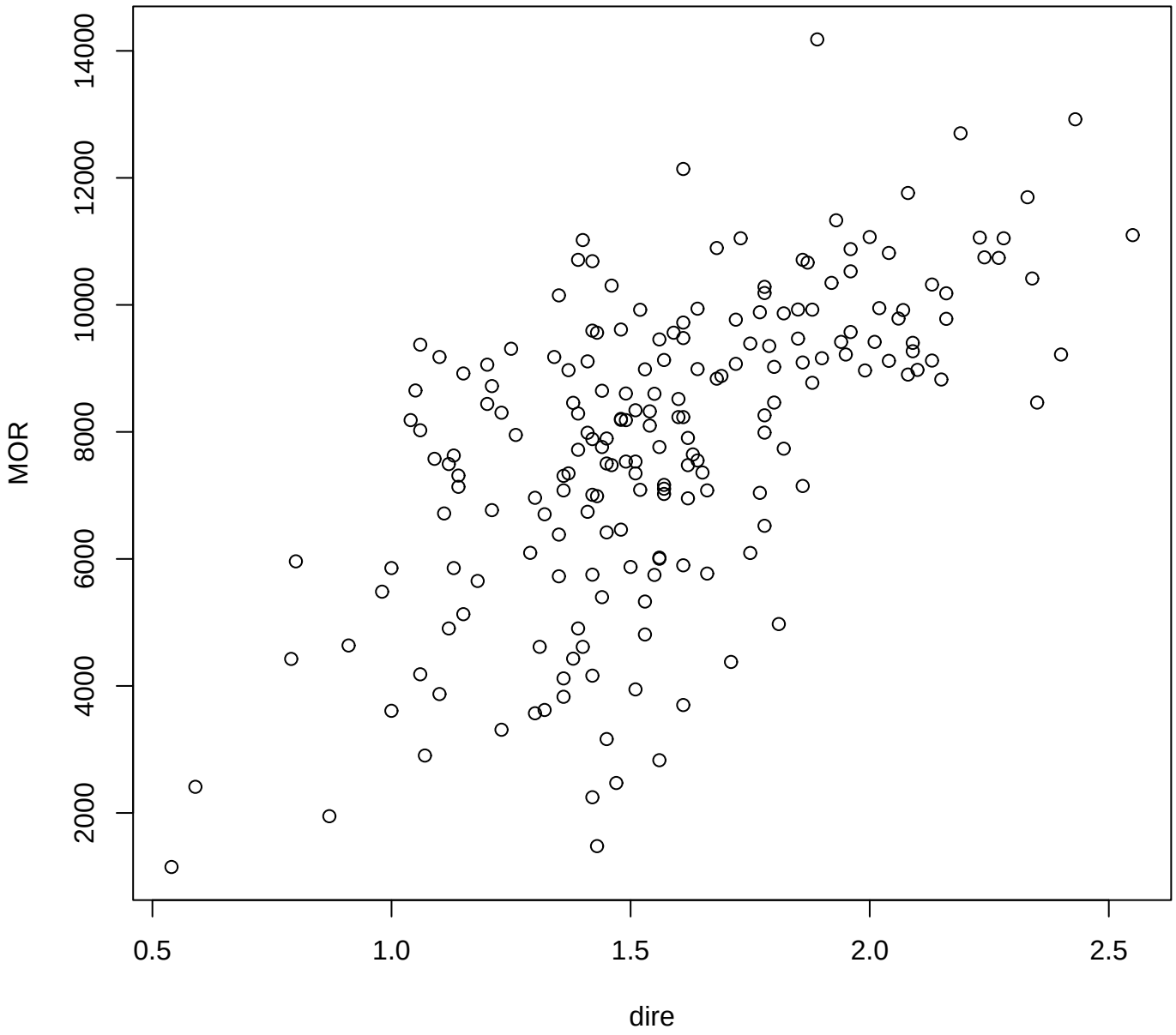


Figure 3: MOR versus dire.

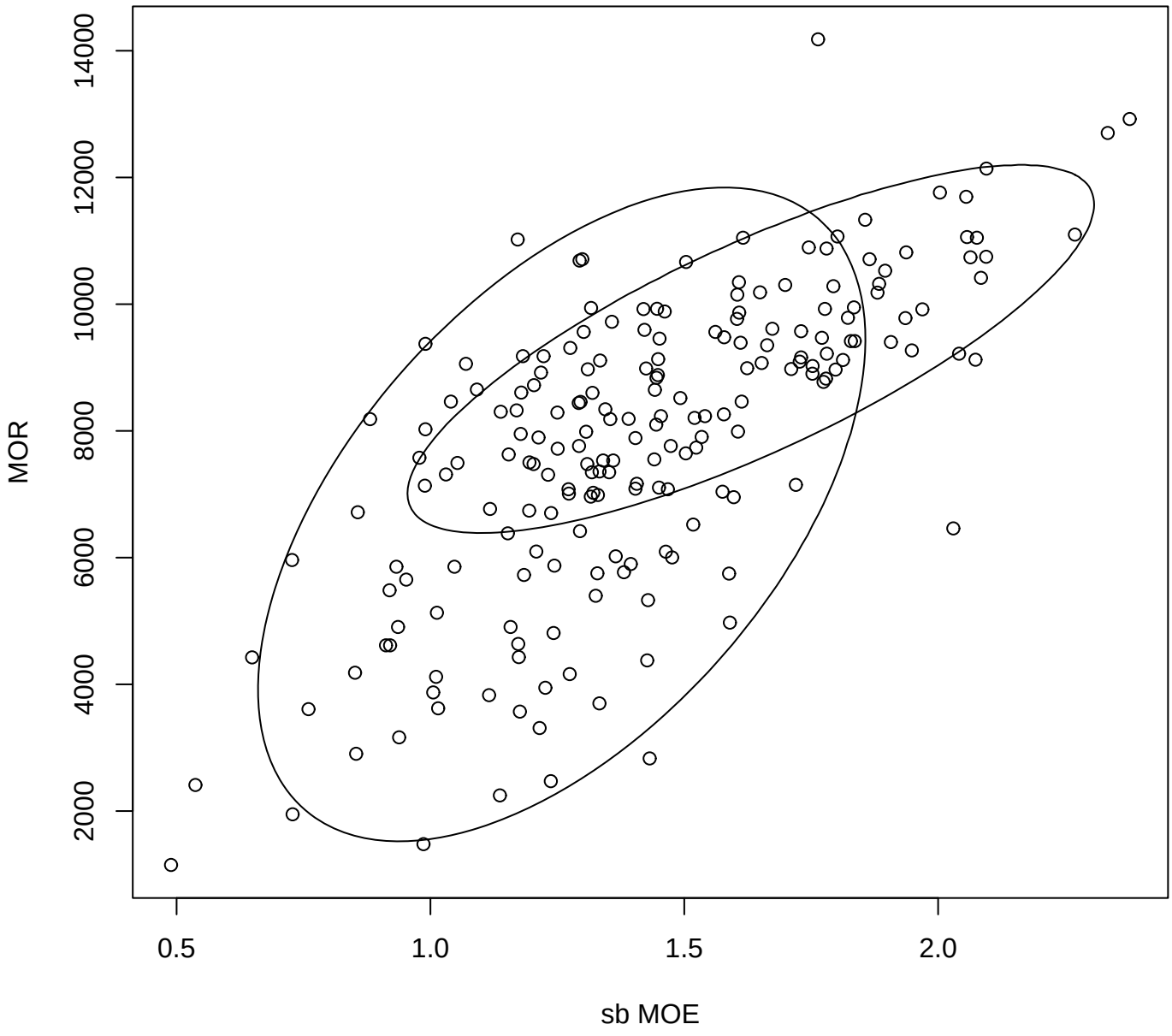


Figure 4: Scatter plot of MOR versus sb MOE, and 0.90 probability content contours for the two bivariate normal components of the fitted mixed bivariate normal model.

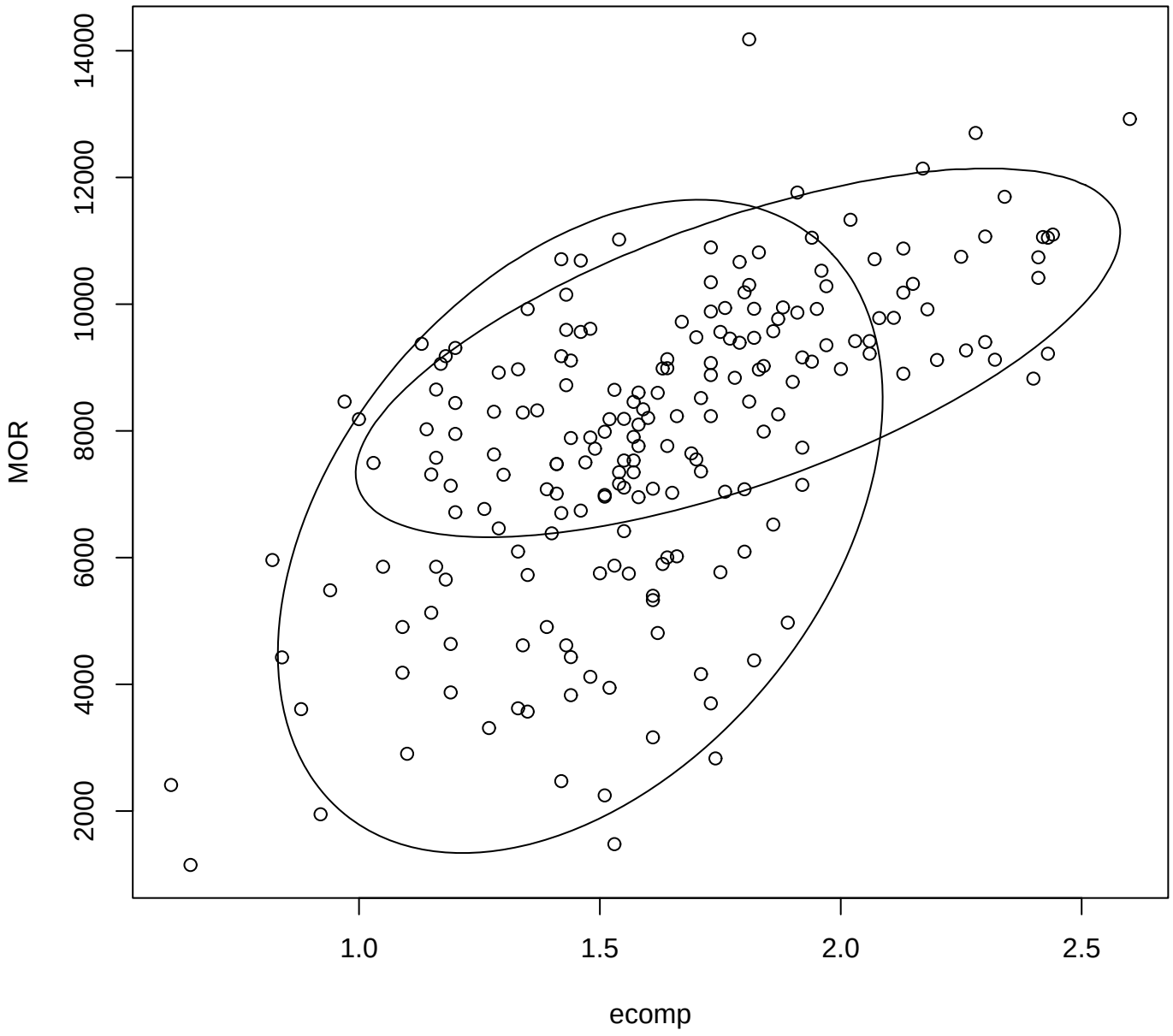


Figure 5: Scatter plot of MOR versus ecomp, and 0.90 probability content contours for the two bivariate normal components of the fitted mixed bivariate normal model.

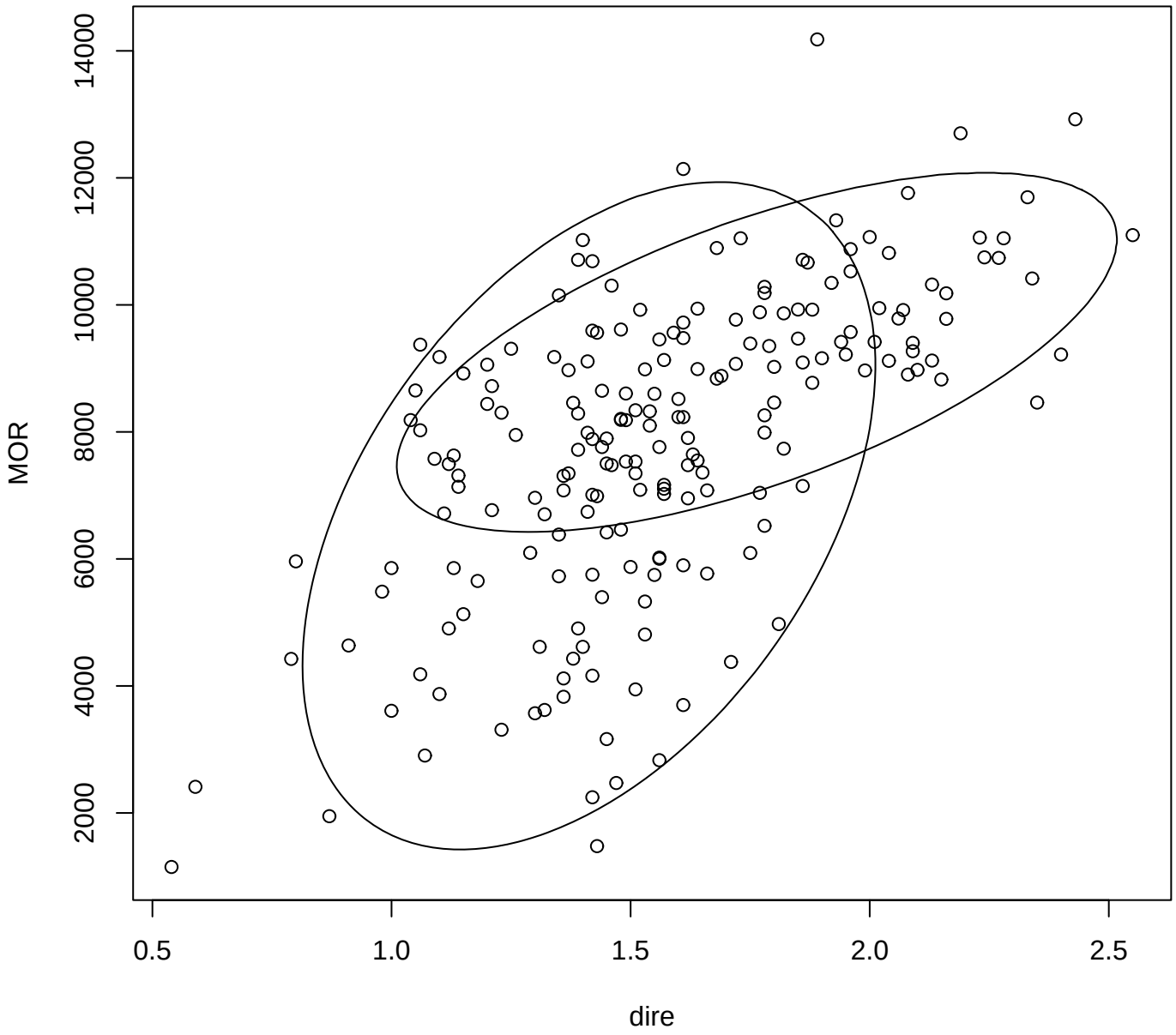


Figure 6: Scatter plot of MOR versus dire, and 0.90 probability content contours for the two bivariate normal components of the fitted mixed bivariate normal model.

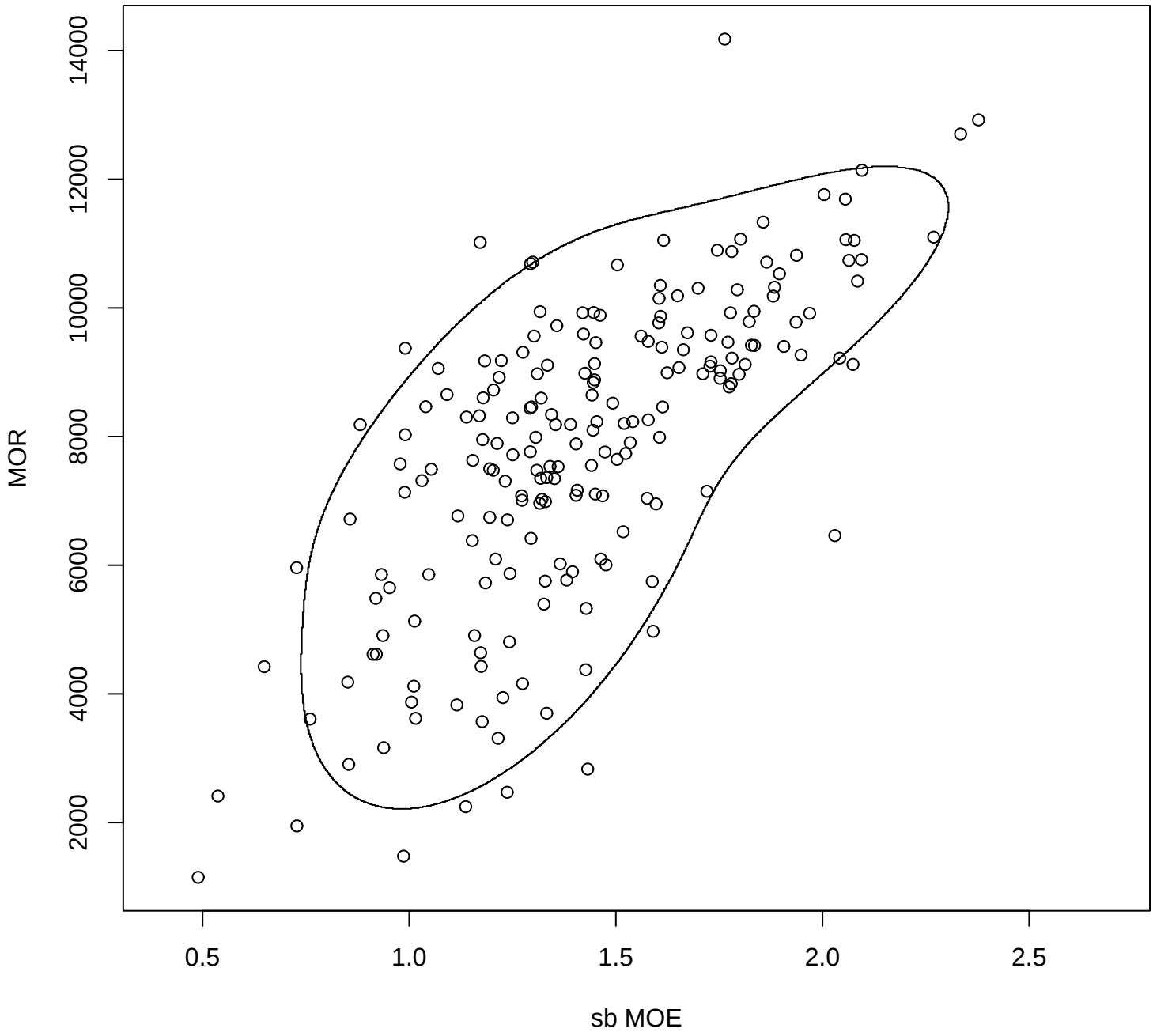


Figure 7: Scatter plot of MOR versus sb MOE, and approximate 0.90 probability content contour for the fitted full mixed bivariate normal model.

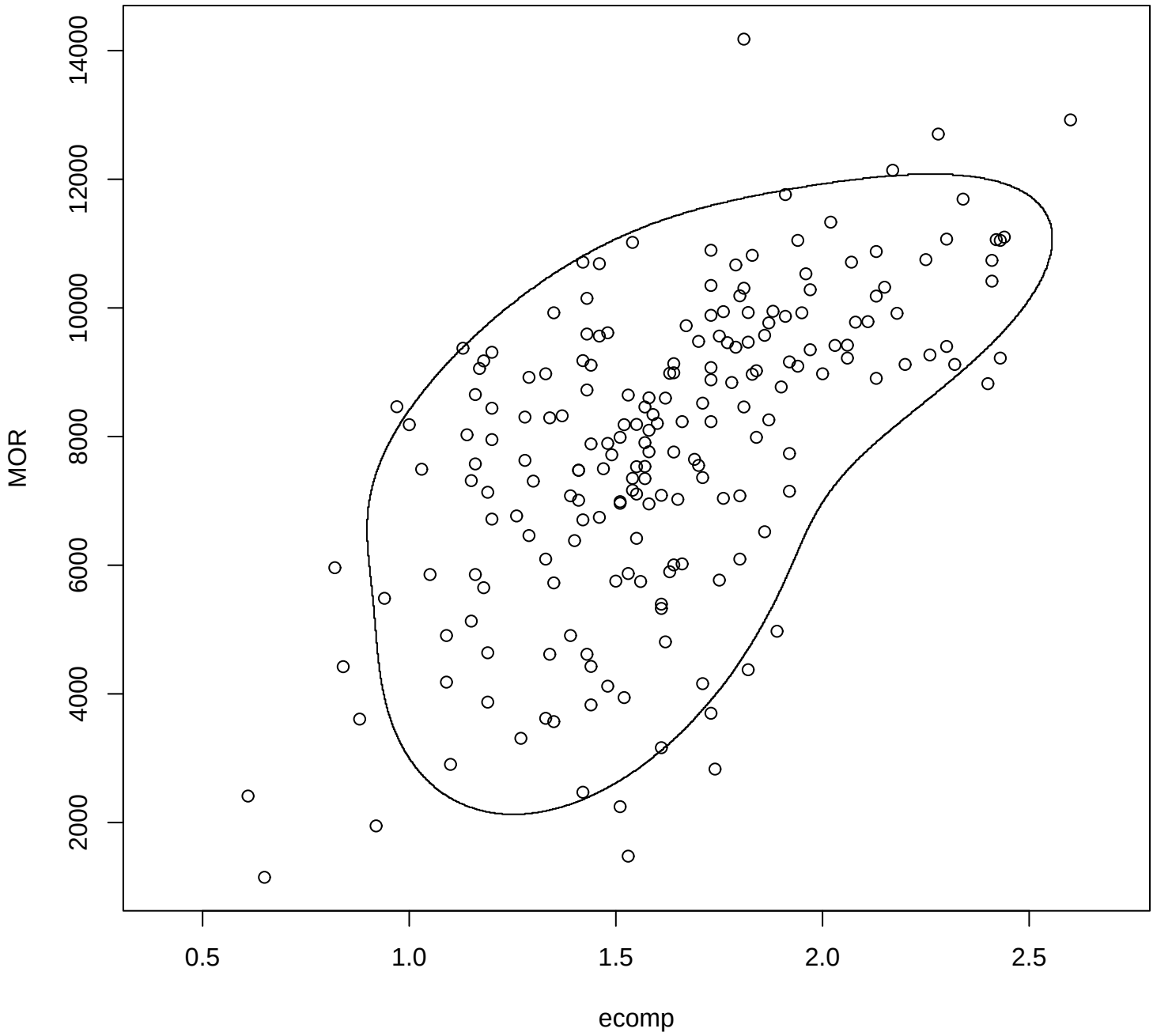


Figure 8: Scatter plot of MOR versus ecomp, and approximate 0.90 probability content contour for the fitted full mixed bivariate normal model.

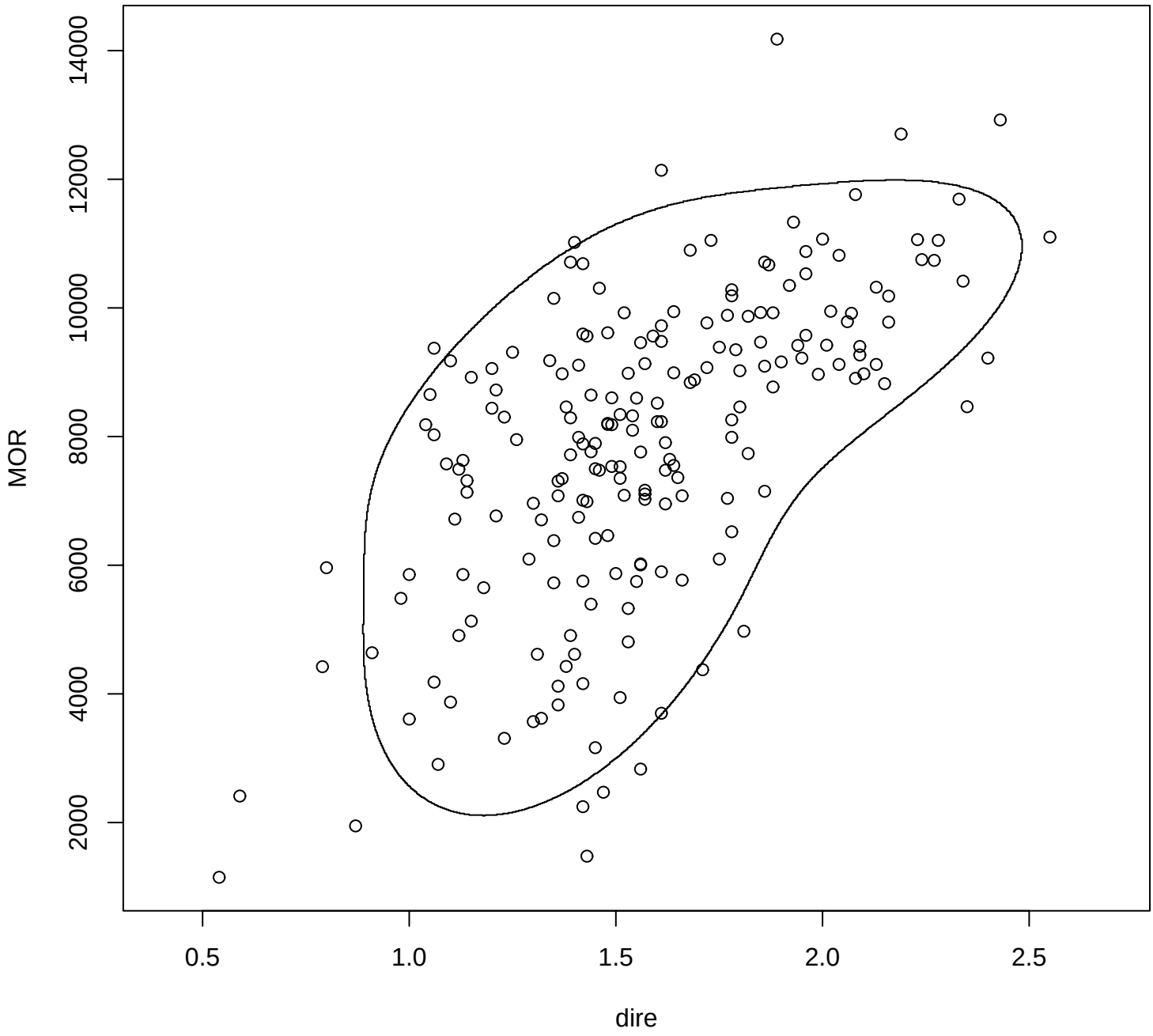


Figure 9: Scatter plot of MOR versus dire, and approximate 0.90 probability content contour for the fitted full mixed bivariate normal model.

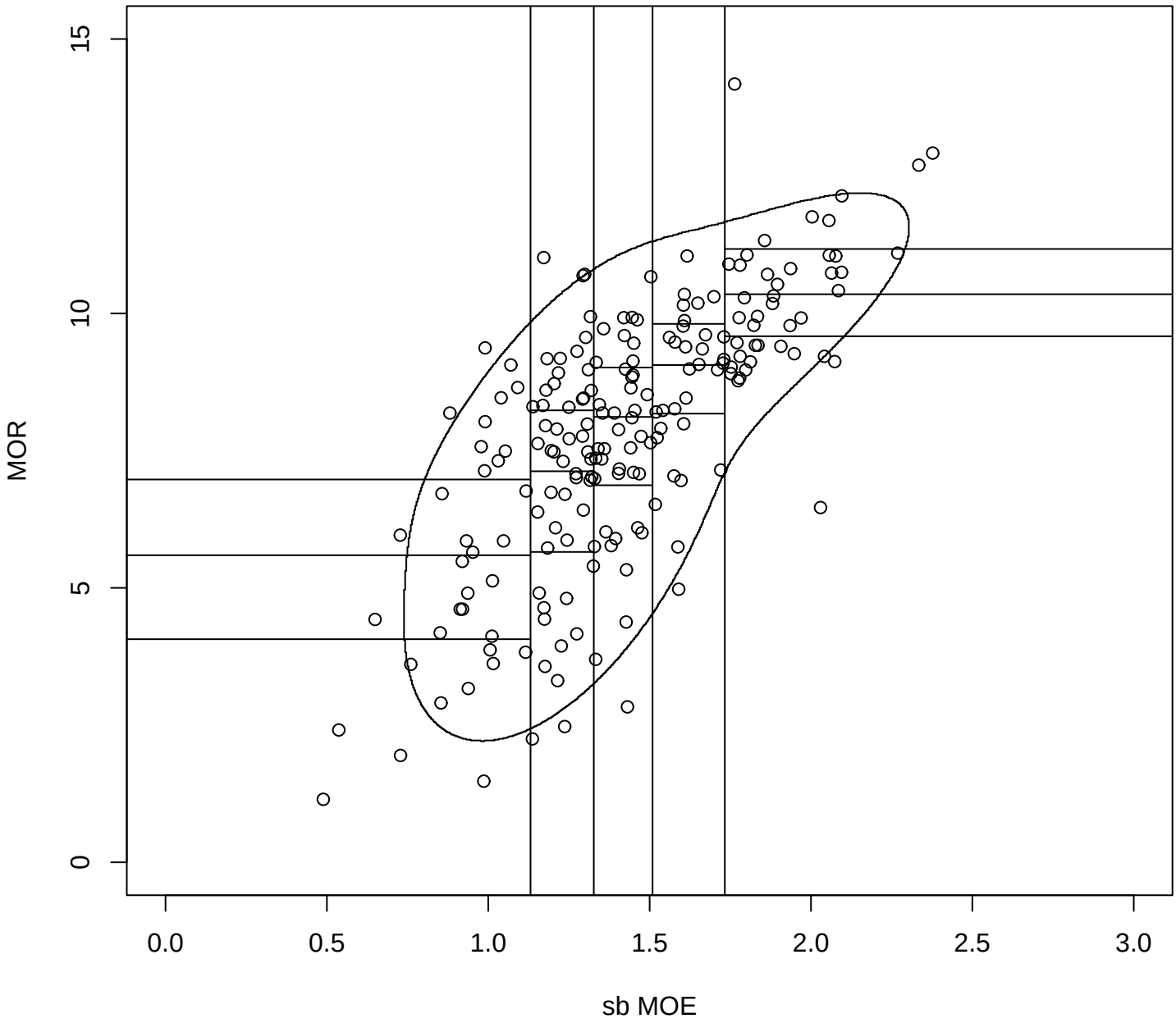


Figure 10: MOR versus sb MOE scatter plot overlaid with the 0.90 probability content contour, and 5×4 "chi-squared goodness-of-fit" cells.

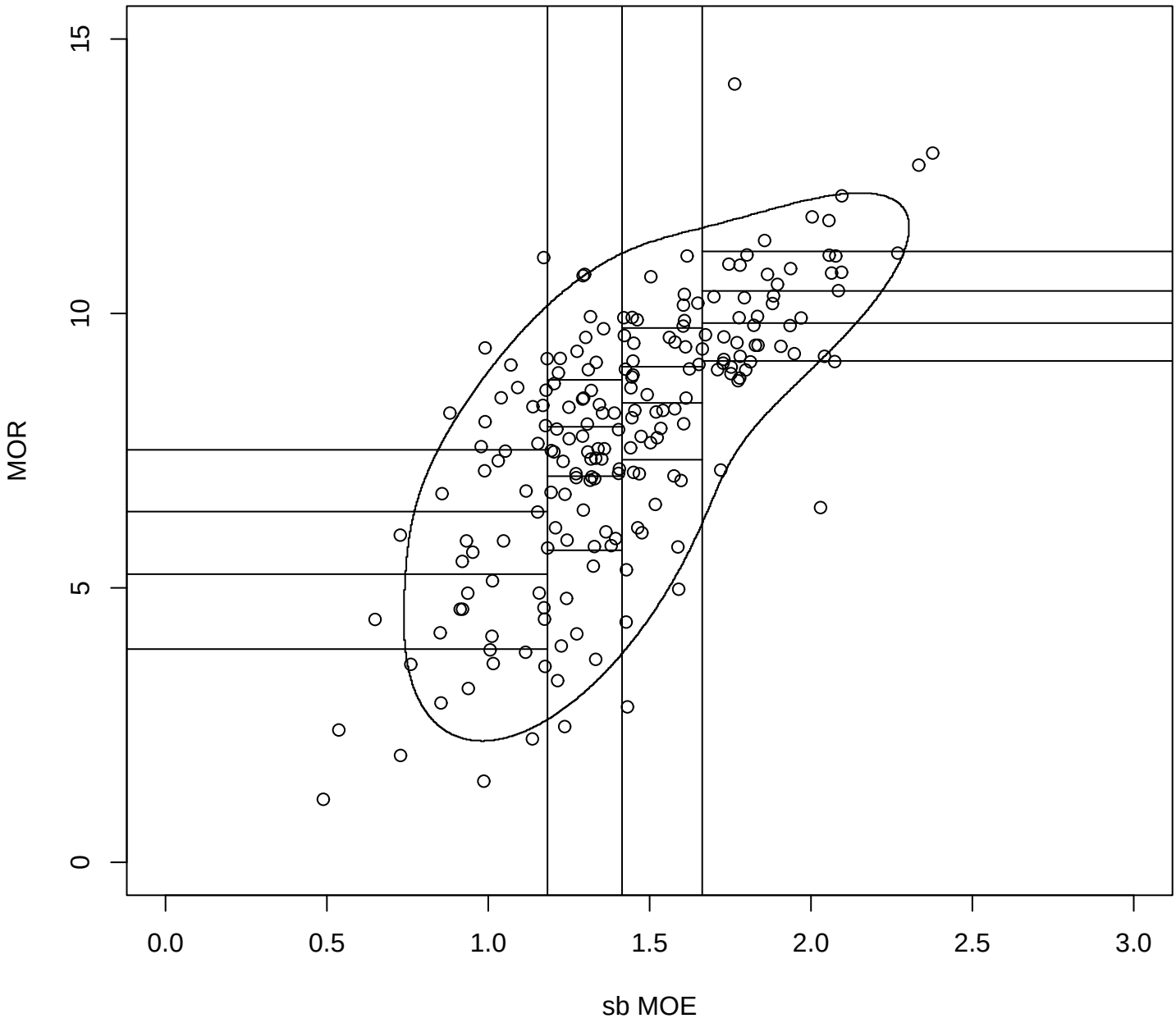


Figure 11: MOR versus sb MOE scatter plot overlaid with the 0.90 probability content contour, and 4×5 "chi-squared goodness-of-fit" cells.

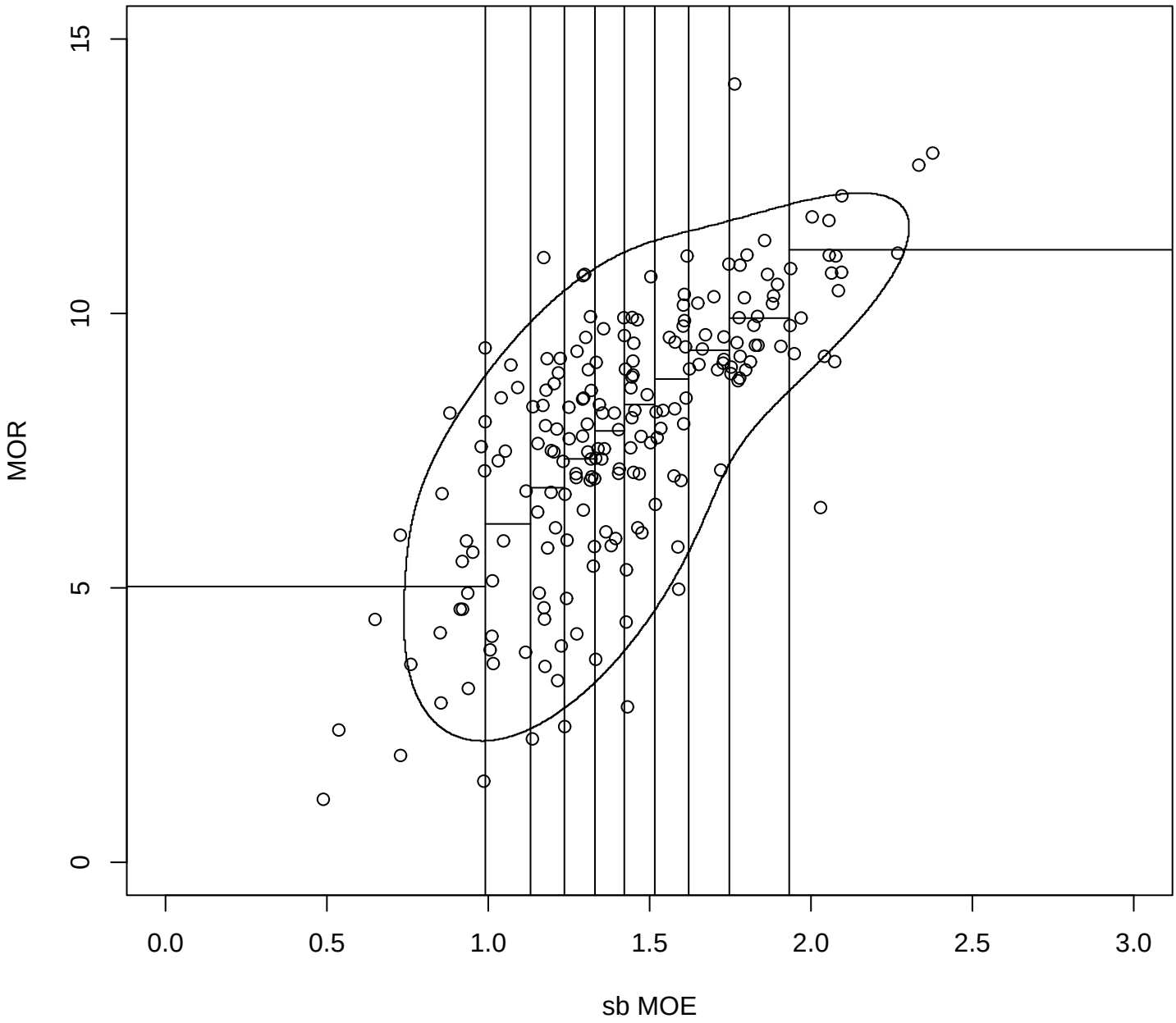


Figure 12: MOR versus sb MOE scatter plot overlaid with the 0.90 probability content contour, and 10×2 "chi-squared goodness-of-fit" cells.

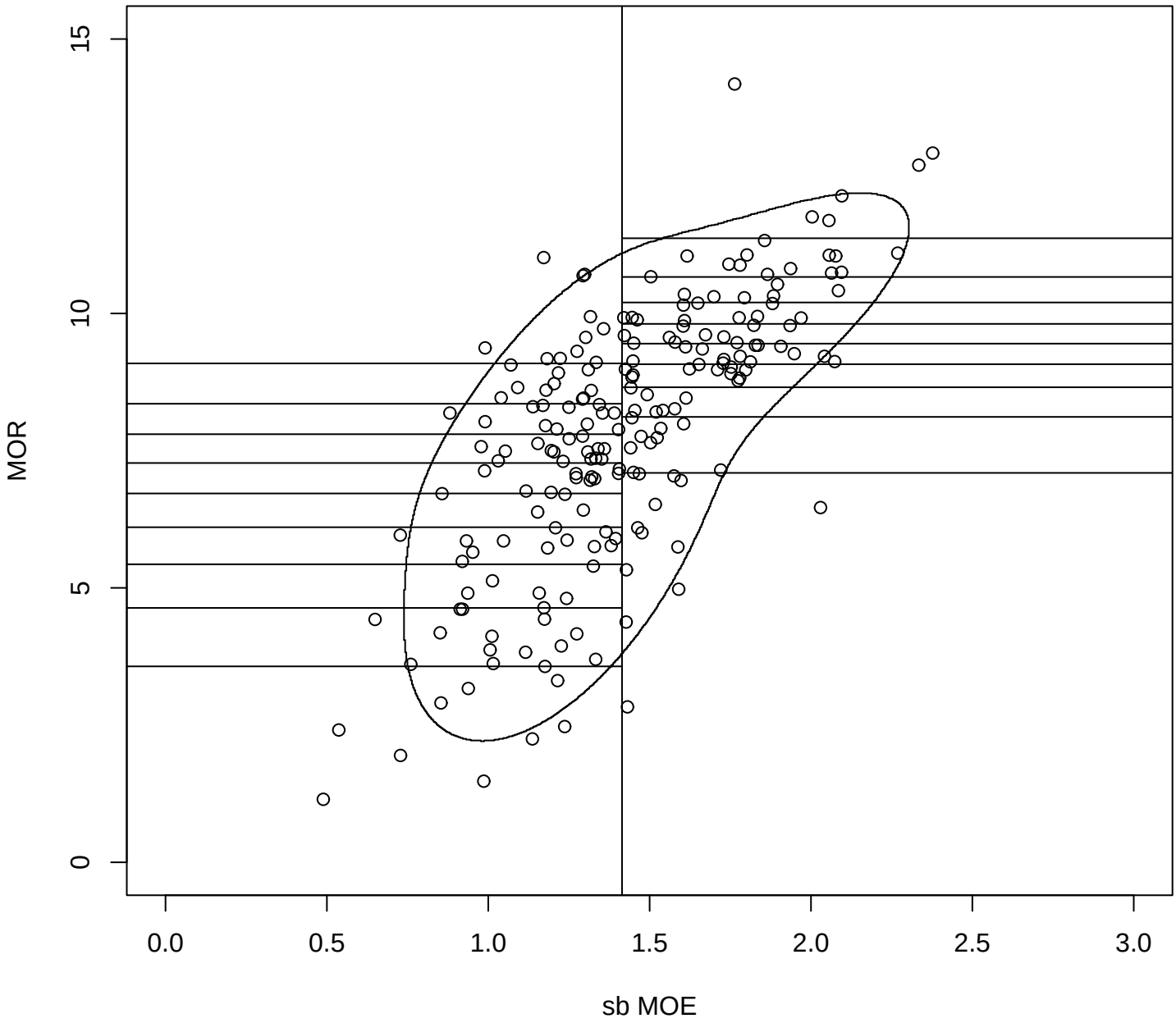


Figure 13: MOR versus sb MOE scatter plot overlaid with the 0.90 probability content contour, and 2×10 "chi-squared goodness-of-fit" cells.

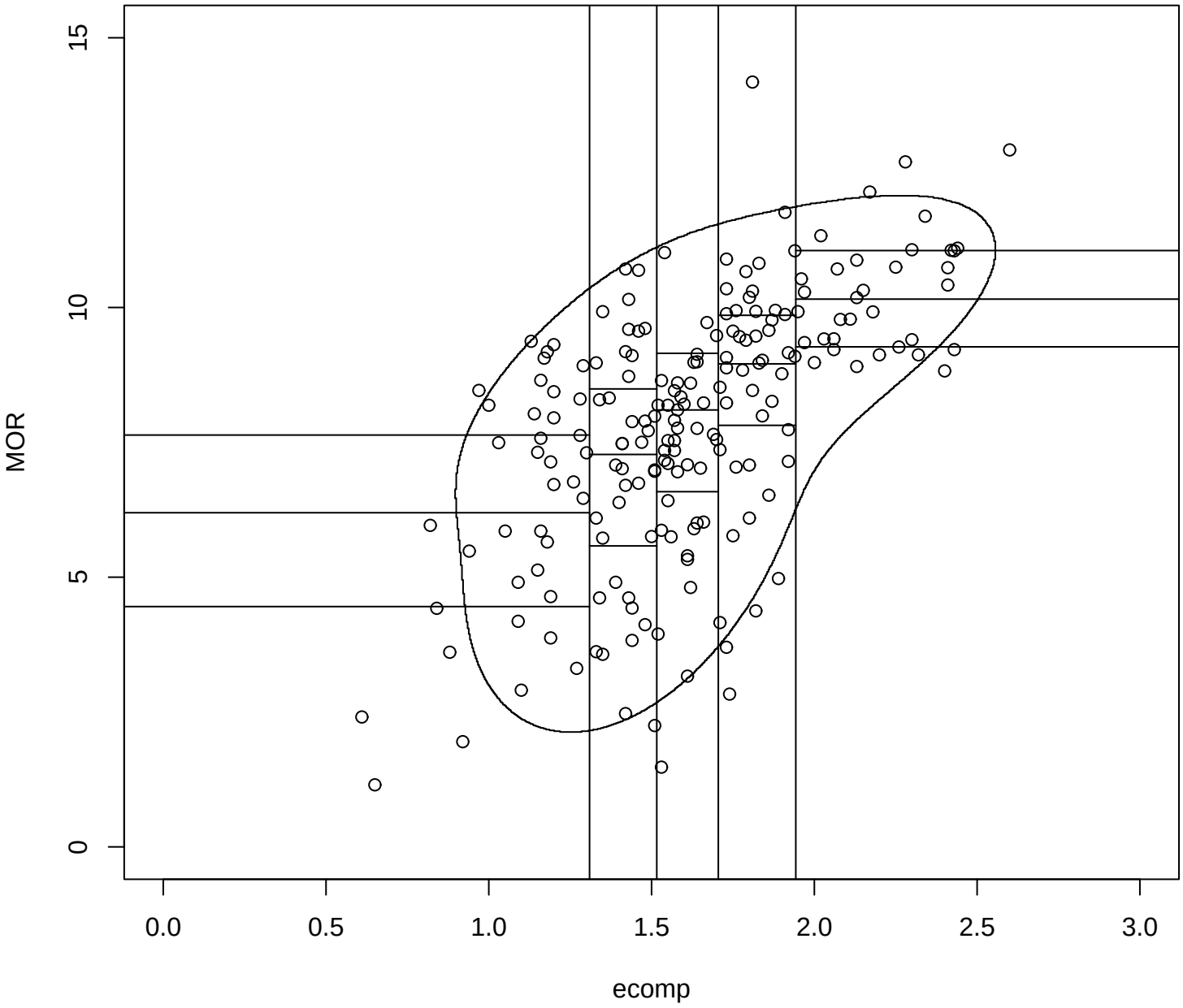


Figure 14: MOR versus ecomp scatter plot overlaid with the 0.90 probability content contour, and 5×4 "chi-squared goodness-of-fit" cells.

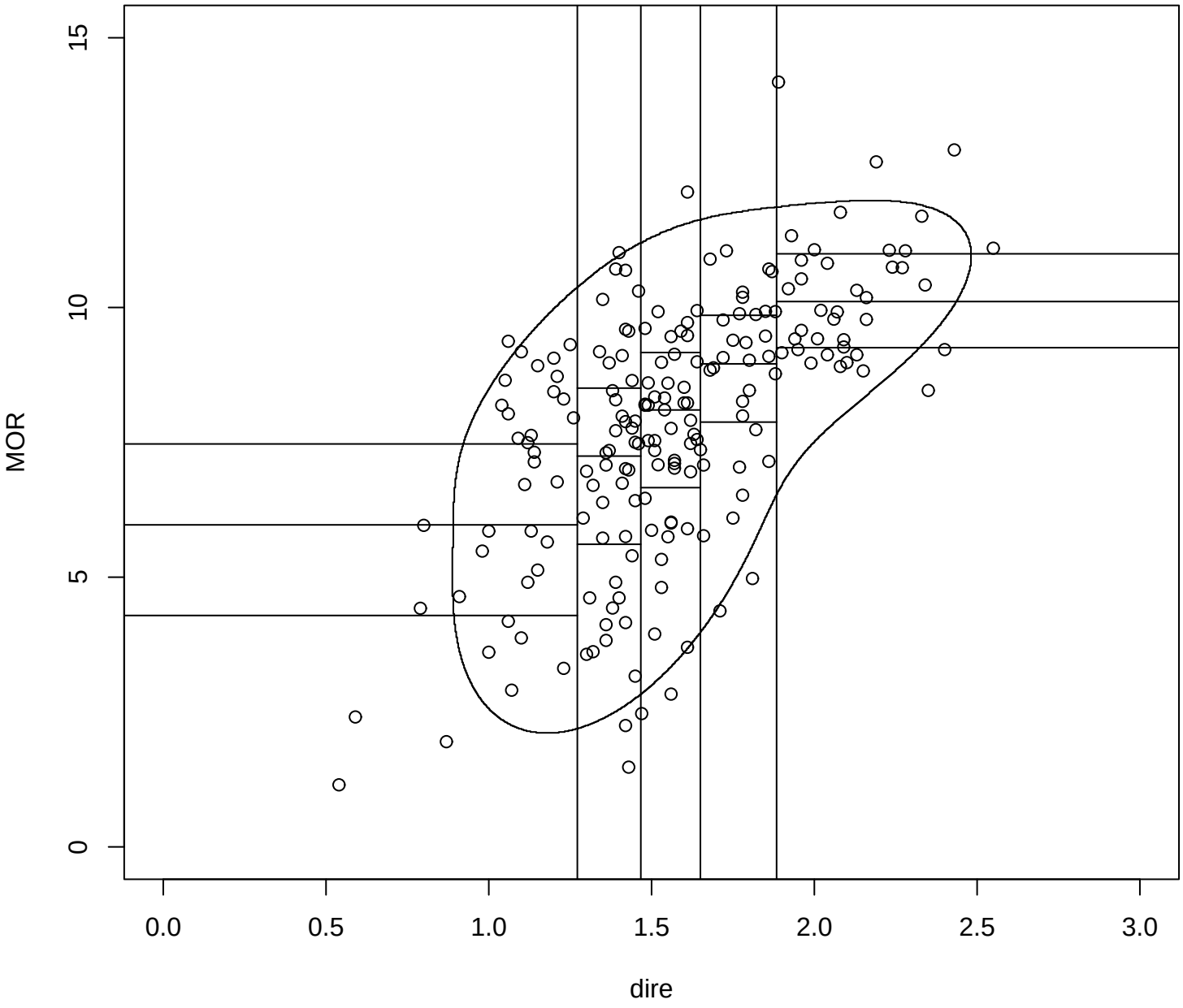


Figure 15: MOR versus dire scatter plot overlaid with the 0.90 probability content contour, and 5×4 "chi-squared goodness-of-fit" cells.

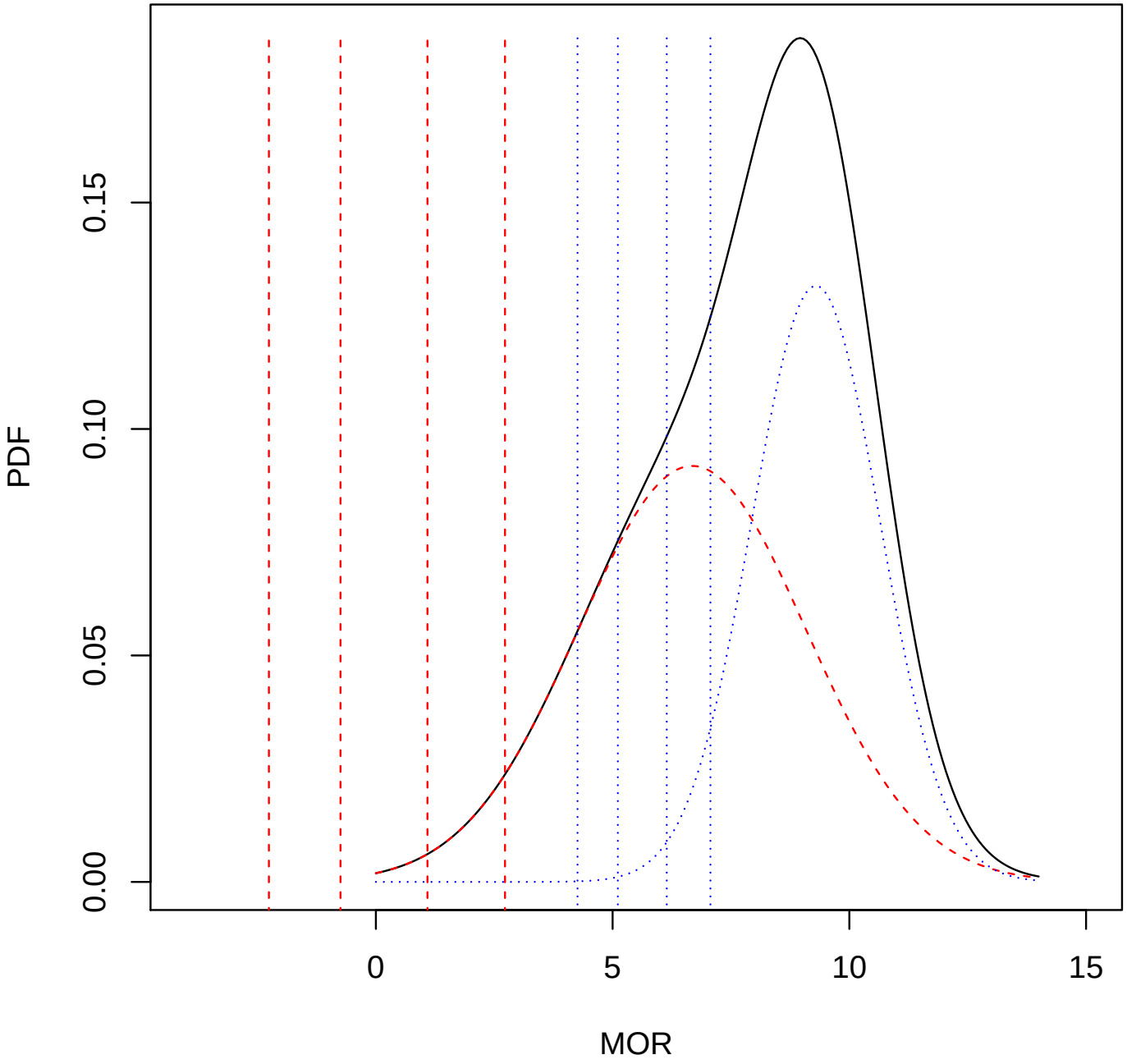


Figure 16: Mixed univariate normal MOR pdf (tall, solid, black) corresponding to the mixed bivariate normal for sb MOE–MOR (prior to pseudo-truncation); left pdf component (dashed red); right pdf component (dotted blue); 0.0001, 0.001, 0.01, 0.05 quantiles for left component (dashed red vertical lines); 0.0001, 0.001, 0.01, 0.05 quantiles for right component (dotted blue vertical lines).

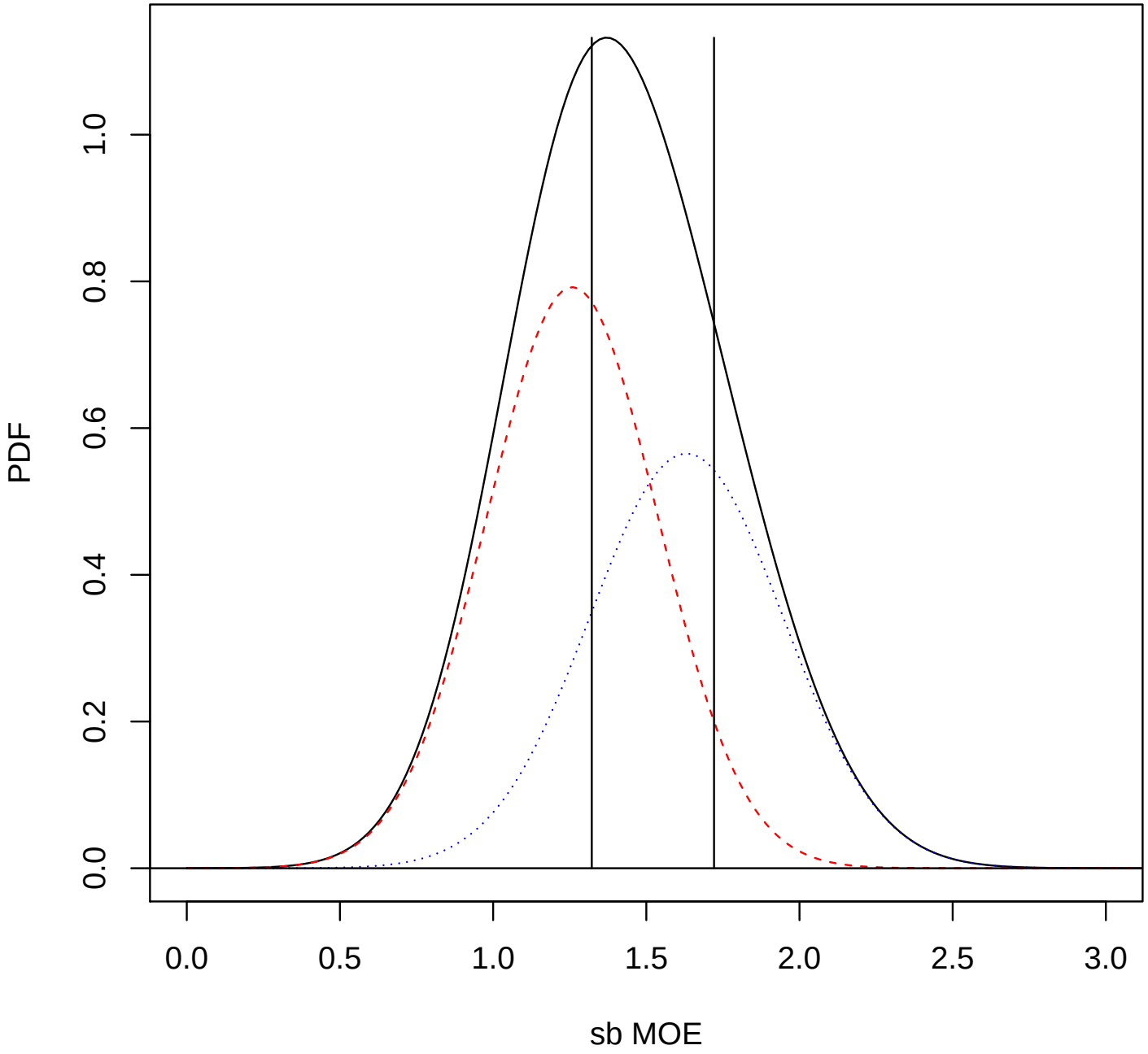


Figure 17: Mixed univariate normal sb MOE pdf (tall, black) corresponding to the mixed bivariate normal for sb MOE–MOR; black vertical lines at 40th and 80th percentiles of the mixed univariate normal sb MOE distribution; left pdf component (dashed red); right pdf component (dotted blue).

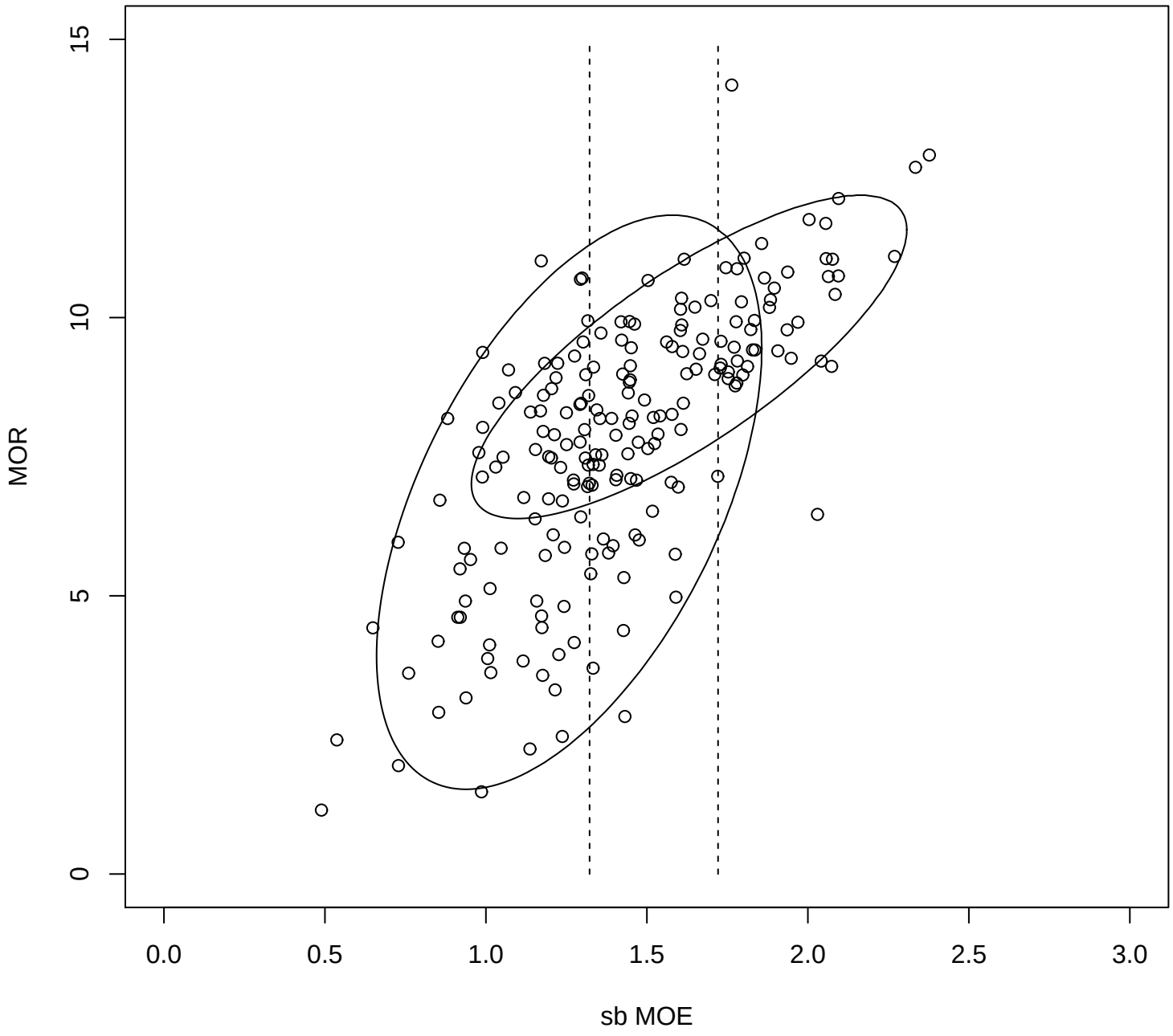


Figure 18: Scatter plot of MOR versus sb MOE, and 0.90 probability content contours for the two bivariate normal components of the fitted mixed bivariate normal model. The vertical lines are at the 0.40 and 0.80 sb MOE quantiles estimated from the marginal sb MOE distribution calculated from the fitted mixed bivariate normal model.

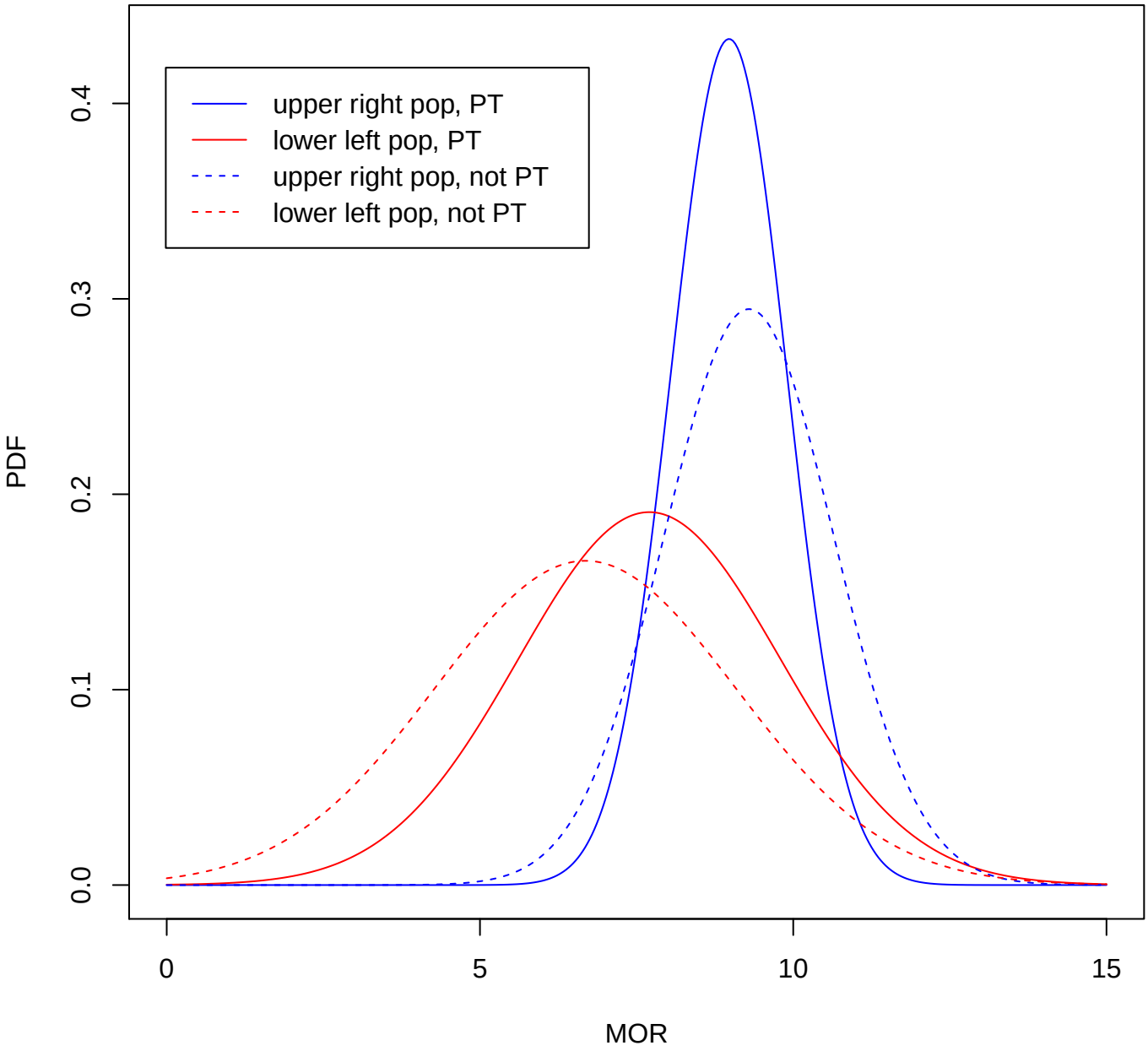


Figure 19: The pdfs of the MORs of the left and right bivariate subpopulations associated with the non-pseudo-truncated mixed bivariate normal model are plotted as dashed lines. The pdfs of the pseudo-truncated (via the 0.40 and 0.80 sb MOE quantiles) MOR distributions associated with the left and right bivariate subpopulations are plotted as solid lines.

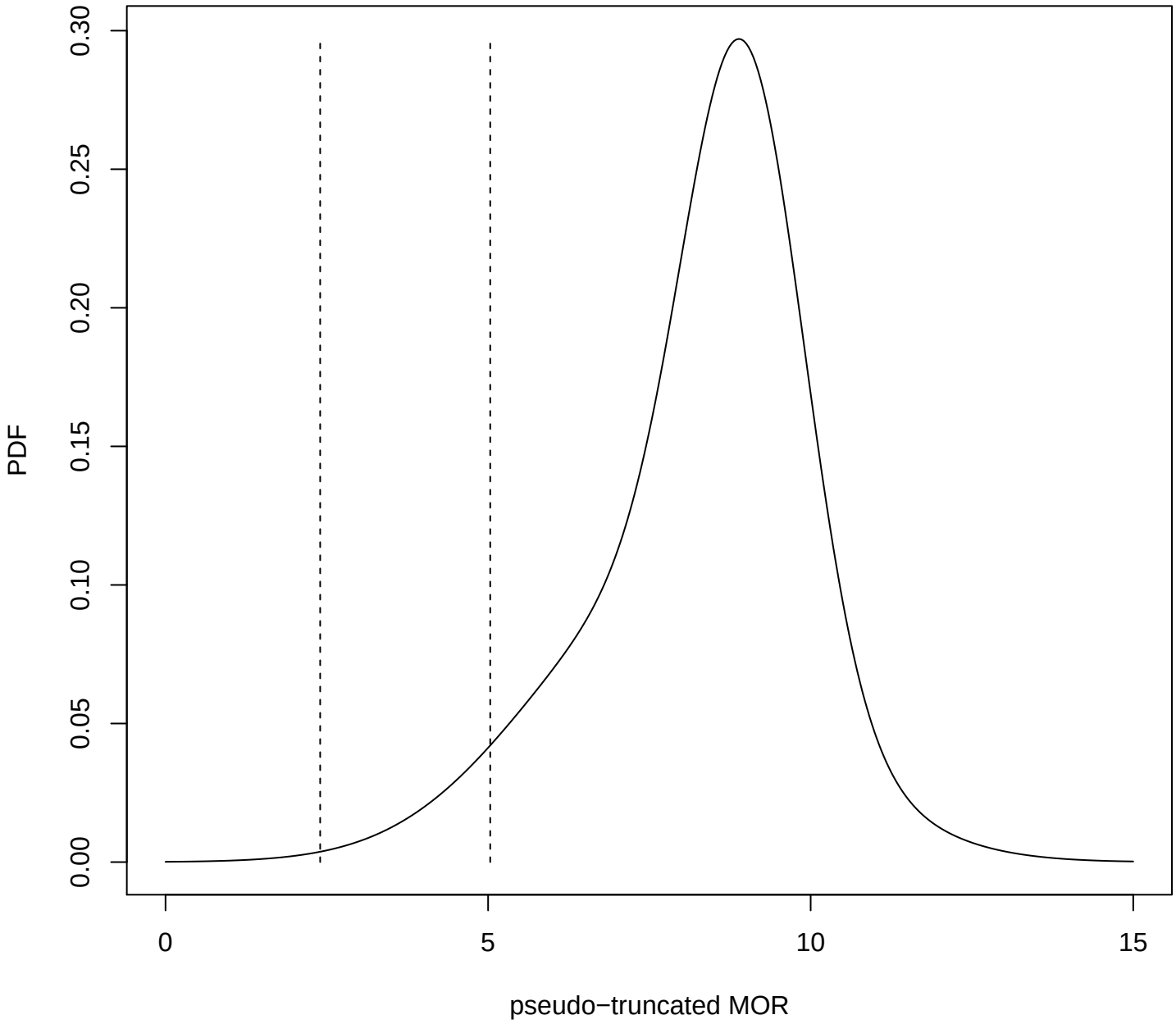


Figure 20: Pseudo-truncated (on sb MOE) univariate mixed normal MOR pdf corresponding to the mixed bivariate normal sb MOE–MOR distribution (for $p = 0.55$). The dashed lines are at the 5th percentile and at the "allowable property" (5th/2.1).

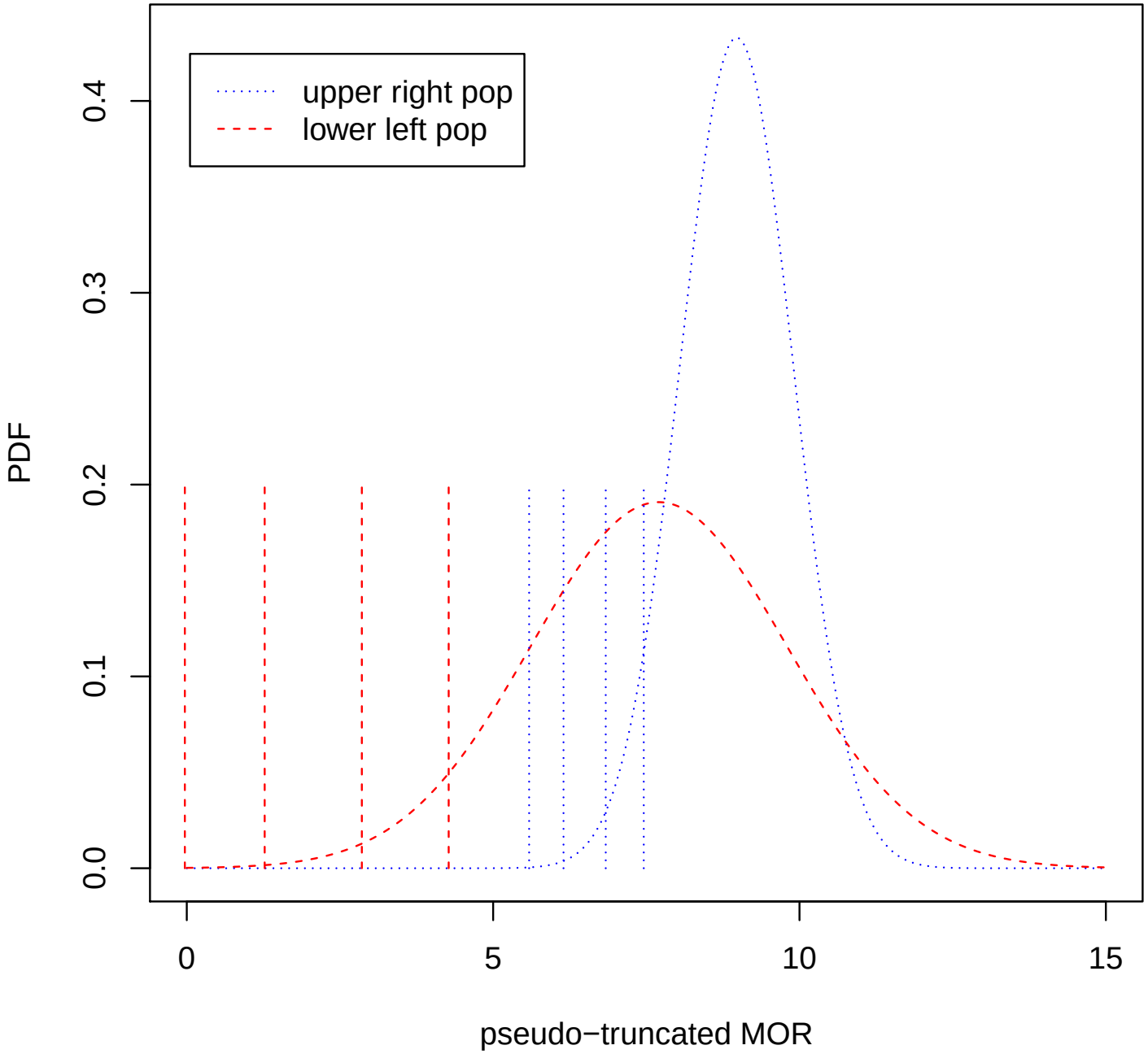


Figure 21: For $p = 1$ (left population, dashed red) and $p = 0$ (right population, dotted blue), we plot the pseudo-truncated (on sb MOE) univariate mixed normal MOR pdfs corresponding to the mixed bivariate normal sb MOE–MOR distribution. The vertical dashed red lines are at the 0.0001, 0.001, 0.01, and 0.05 quantiles of the left population. The vertical dotted blue lines are at the 0.0001, 0.001, 0.01, and 0.05 quantiles of the right population.