Published in final edited form as:

Teach Stat. 2013 ; 35(3): 131–139. doi:10.1111/test.12012.

Technology-enhanced Interactive Teaching of Marginal, Joint and Conditional Probabilities: The Special Case of Bivariate Normal Distribution

Ivo D. Dinov, Scott Kamino, Bilal Bhakhrani, and Nicolas Christou

Statistics Online Computational Resource University of California, Los Angeles, Los Angeles, CA 90095, USA

Summary

Data analysis requires subtle probability reasoning to answer questions like What is the chance of event A occurring, given that event B was observed? This generic question arises in discussions of many intriguing scientific questions such as What is the probability that an adolescent weighs between 120 and 140 pounds given that they are of average height? and What is the probability of (monetary) inflation exceeding 4% and housing price index below 110? To address such problems, learning some applied, theoretical or cross-disciplinary probability concepts is necessary. Teaching such courses can be improved by utilizing modern information technology resources. Students' understanding of multivariate distributions, conditional probabilities, correlation and causation can be significantly strengthened by employing interactive web-based science educational resources. Independent of the type of a probability course (e.g. majors, minors or service probability course, rigorous measure-theoretic, applied or statistics course) student motivation, learning experiences and knowledge retention may be enhanced by blending modern technological tools within the classical conceptual pedagogical models.

We have designed, implemented and disseminated a portable open-source web-application for teaching multivariate distributions, marginal, joint and conditional probabilities using the special case of bivariate Normal distribution. A real adolescent height and weight dataset is used to demonstrate the classroom utilization of the new web-application to address problems of parameter estimation, univariate and multivariate inference.

© 2013 The Authors Teaching Statistics© 2013 Teaching Statistics Trust

Dinov@ucla.edu.

Changes to Teaching Statistics

We now have an online system for article submission. Please visit www.teachingstatisticsonline.org for more information. Whilst implementing this system we have also created two new article types: "Classroom notes" and "Assessment matters". The aim of the former is that we have single paragraph suggestions for classroom activity, with perhaps only a single reference to prior art. We do not intend publishing these as separate articles, but would like an article in every issue to contain a number of "bite size ideas" that classroom practitioners can use with minimal modification.

"Assessment matters" reflects the importance of designing assessment in order that appropriate learning takes place.

We would also be extremely grateful if potential reviewers registered on the site. One of our goals is to speed up the review process; and the more closely we can match articles with reviewer interests hopefully the better the whole process becomes. Thanks again to the many reviewers who help in the production of this journal.

Keywords

teaching; science education; blended instruction; statistics education; webapp; technology-enhanced

INTRODUCTION

Diverse arrays of interesting scientific questions that frequently arise in everyday life may require probability reasoning to obtain exact or approximate answers. Examples of such inquiries include Given that an adolescent is of average height, what is the probability the adolescent weighs between 120 and 140 pounds?, What is the probability of (monetary) inflation exceeding 4% and housing price index below 110? and What is the probability that the second card drawn (without replacement) from a standard 52-card deck is a diamond given that the first drawn card was a king? These questions provide specific examples of conditional, joint and marginal distributions in multivariate (bivariate, in this case) settings, respectively. Generally, there are two strategies for solving such problems – analytical (using theory and first-principles) or experimental (using simulations, approximations or estimations). Solutions to such challenges may be specific or generic, numeric or qualitative, algorithmic or computational.

In multidisciplinary probability courses, utilizing modern information technology (IT) resources (Algozzine et al. 2011; Huang et al. 2011) may impact the instructors' teaching strategies and enhance students' learning and classroom experiences. Understanding joint, marginal and conditional probabilities for modeling and comprehension of the world around us requires both rigorous training and at the same time effective pedagogical approaches for teaching visual, active or didactic learners (Dietz and Stevenson 2011; Jahnavi and Patra 2012). The basic ideas of multivariate distributions, conditional probabilities and correlation may lead to misconceptions or fallacies in some students' understanding that can be avoided by employing interactive web-based science education resources (Dinov and Christou 2011; Hwang et al. 2011). Building on the available Statistics Online Computational Resource (SOCR) infrastructure (Dinov 2006; Christou and Dinov 2011), we demonstrate one example of an IT-enhanced interactive instructional resource for teaching multivariate distributions, marginal, joint and conditional probabilities using the special case of the bivariate Normal distribution.

BIVARIATE PROBABILITIES

Suppose we are observing a natural process that generates two (or more) observations X and Y . For example, earthquakes' location, magnitude and depth (Ayalew et al. 2011); vital signs of humans (body temperature, arterial pressure and respiration) (Henneman et al. 2012); stellar mass, radius and luminosity (Newton et al. 2011); phenotypic (age, race and morbidity) and genotypic characteristics (single nucleotide polymorphisms) in childhood-onset schizophrenia (Stephens et al. 2012), and so on. Such random observations of continuous or discrete processes may be thought of as instance measurements. The bivariate Normal distribution has some interesting properties (Lancaster 1957; Fraser and Streit 1980; Patel and Read 1996). The density function of bivariate Normal distribution is defined in

terms of a two-dimensional mean–vector, specifying the center location of the distribution in the two-dimensional plane, and a 2×2 variance–covariance matrix, which characterizes the amount of dispersion along the main X and Y directions, as well as the correlation between X and Y . For the bivariate Normal distribution, the two marginal distributions are also (univariate) normal with means and variances according to the corresponding components of the bivariate distribution parameters (mean–vector and the variance–covariance matrix). Finally, the conditional probability distributions are also normal with explicit closed form expressions for their mean and dispersion parameters in terms of the corresponding bivariate density. The Appendix section includes the conceptual definitions and theoretical formulations necessary to apply and compute marginal, conditional and joint probability calculations, as well as the explicit expressions for the marginal and conditional probabilities.

Webapp

To improve students' learning experiences and augment the instructors' strategies of teaching probability theory and multivariate distributions, we designed, implemented and openly disseminated an interactive web-application (webapp) that demonstrates these concepts using bivariate Normal distribution (Castillo and Galambos 1989). The SOCR (Dinov and Christou 2009) bivariate Normal distribution webapp (<http://socr.ucla.edu/htmls/HTML5/BivariateNormal/>) requires a modern web-browser with HTML and JavaScript support. It includes a three-dimensional view of the bivariate distribution (based on WebGL) and includes a paired activity (see below). Instructors and learners may use the *settings* panel to initialize the webapp. The *control* panel allows selection of the appropriate bivariate limits for the X and Y variables and specification of the desired joint, marginal or conditional probability function. In addition to the interactive three-dimensional graph for the joint probability, the webapp provides a complementary one-dimensional Normal distribution graph for computing user-specified marginal or conditional distributions. Entering any non-numerical characters in the settings or control panels resets the parameter to infinity (∞).

The webapp allows three-dimensional rotation of the bivariate Normal distribution by clicking and dragging the mouse on the graph. For browsers without WebGL support, the three-dimensional graph shows a wireframe distribution instead of the standard triangulated surface. The webapp shows all probability results a text area. The webapp, Figure 1, is HTML5-standard compliant and should work on all modern browsers (including mobile devices) without any special installations or configurations.

LEARNING ACTIVITY

We designed a complementary hands-on learning activity (http://wiki.stat.ucla.edu/socr/index.php/SOCR_BivariateNormal_JS_Activity) that aims the following:

- Provide a visualization tool for better understanding of multivariate distributions (e.g. bivariate Normal distribution).
- Clarify the definitions and interplay between marginal, conditional and joint probability distributions.

- Demonstrate how to calculate the marginal, conditional and joint probabilities for Bivariate Normal.
- Show how correlation influences the distribution of two normally distributed variables. Demonstrate that when X and Y have joint bivariate Normal distribution with zero correlation, then X and Y must be independent.
- Build a framework for generalizing the univariate Normal distribution to higher dimensions.

In our experience, teaching probability theory and multivariate distributions, students have responded overwhelmingly positively to the experience of interactively computing marginal, conditional and joint bivariate (Normal) probabilities, formulating research hypotheses, quickly validating them using the webapp and answering concrete scientific questions using the univariate and bivariate Normal distribution calculators. Indeed, instructors embedding the bivariate Normal distribution webapp and activity in their curricula need to demonstrate in class the tight integration of the course-specific conceptual materials, the applications they are addressing, and the graphing and computing resources provided in the webapp.

APPLICATION: HUMAN HEIGHTS AND WEIGHTS

One example demonstrating the classroom utilization of the new bivariate Normal distribution webapp addresses the first specific scientific question we presented in the Introduction: Given that an adolescent is of average height, what is the probability the adolescent weighs between 120 and 140 pounds? To answer this question, we first start with some real data about adolescents' Heights (measured in inches, *in*) and Weights (measured in pounds, *lb*) (http://wiki.stat.ucla.edu/socr/index.php/SOCR_Adolescents_HeightsWeights) (Leung et al. 1996; So et al. 2008). All the calculations and interpretations below explicitly assume that the *data* are representative of the population of adolescents, which is the focus of the driving motivation questions. If this assumption is invalid, the statistical inference may be misguided. When there is a discrepancy between the population we are trying to make inference about and the data available for quantitative analysis, one needs to redesign the experiment starting with the initial data collection to ensure the data represent a *random sample* from the target population.

Parameter estimation

We can start by copy-pasting the data (initial 200 observations, or even the entire 25,000 row dataset) in the SOCR Modeler (http://socr.ucla.edu/htmls/SOCR_Modeler.html). Then, we can estimate the *mean and standard deviation* of the Height and Weight variables, as well as investigate the distributions of the two variables. By looking at these distributions, we can conclude that Height and Weight roughly follow a Normal distribution, Figure 2. Quantile–Quantile probability plots (Al-Aziz et al. 2010; Dinov and Christou 2011) and Kolmogorov–Smirnov test (Che et al. 2009) can be employed to obtain additional qualitative and quantitative support for normality of the weight and height measurements in this large sample. The SOCR Modeler generates estimates of the mean height (67.95 in), height standard deviation (1.94 in), mean weight (127.22 lb) and the standard deviation of

weight (11.96 lb), Figure 3. The results also indicate the strength of the statistical evidence to reject a null hypothesis that the sample distribution may be modeled by Normal distribution (with the estimated mean and standard deviation).

To estimate the *correlation* between height and weight, we use the SOCR linear regression applet (<http://www.socr.ucla.edu/htmls/ana/SimpleRegressionAnalysis.html>). Mapping height as dependent and weight as independent variables, and clicking ‘Calculate’ computes a linear model including the correlation between height and weight ($\hat{\rho}=0.557$), Figure 4.

So far, we have X = height, Y = weight variables with the following distributions:

$$X \sim N(\hat{\mu}_X=67.95, \hat{\sigma}_X^2=1.94^2), Y \sim N(\hat{\mu}_Y=127.22, \hat{\sigma}_Y^2=11.96^2) \text{ and}$$

$\hat{\rho}=\text{Corr}(X, Y)=0.557$. Now, plugging these five estimated quantities (the *hat* notation ‘^’ indicates data-driven parameter estimates) into the SOCR Bivariate Normal webapp (<http://socr.ucla.edu/htmls/HTML5/BivariateNormal/>) allows us to explore the marginal, joint and conditional probabilities.

Marginal probabilities

When looking at marginal distributions, we disregard the variable(s) we do not care about. For the bivariate normal case, the marginal distribution of a single variable is the (univariate) distribution of the variable itself.

Disregarding weight, what is the probability that a randomly selected adolescent will be between 65 and 70 inches tall? Look at the marginal distribution of height. The probability that the height of (a randomly selected) adolescent will be between 65 and 70 in is $P(65 \leq X \leq 70) = 0.799$, Figure 5.

The histogram for weight spanned an interval of 92–165 lb, Figure 2, and the marginal (model) distribution of weight is supported on the interval [79, 175] pounds. This interval is larger because the (sample) histogram of weight is not perfectly symmetric. The Normal distribution model does a good job of capturing the overall shape of weight (see above). The histogram of height spans an interval from 62 to 75 in, Figure 2, and the support of the (model) marginal distribution of height in the webapp is [60, 75] inches. Indirectly, this agreement is largely because the distribution of human heights is known to be symmetric, bell-shaped and unimodal (Gudbjartsson et al. 2008; Weedon et al. 2008) and can be approximated well by a Normal distribution. Specifying desired X and Y limits in the Controls panel allows for the calculations of concrete marginal probabilities.

Joint probability

What is the probability of a randomly selected adolescent being between 65 and 70 inches tall and weighing between 120 and 140 pounds? To answer this question, change the Bivariate Normal limits to $120 < X < 140$ and $65 < Y < 70$. Figure 6 shows the graphical illustration of this joint probability and the corresponding quantitative results.

Conditional probability

The answer to one of the original questions, $P(120 \leq Y \leq 140|X = 67.95)$, Given that an adolescent is of average height, what is the probability the adolescent weighs between 120 and 140 pounds?, requires the conditional distribution of weight given a known height. First, we need to change the Bivariate Limits so height only includes the specified average height: 67.95. Then we can compute the conditional probability that an adolescent weighs between 120 and 140 lb, given they are of average height, $P(120 \leq Y \leq 140|X = 67.95) = 0.661$, Figure 7. Note that the last conditional distribution is univariate (in Y). Thus, to generate the three-dimensional plot on the top-left of Figure 7, we extended the conditional range of X to $66.95 < X < 67.95$.

DISCUSSION

Curricular integration of technology, observational and research-derived data, and interactive simulations has potential to improve K-16 science education and enhance the learning experiences of students at different levels (Jordan and Haines 2006; Franklin et al. 2007; Holbrook and Rannikmae 2007; Shelley et al. 2009). The availability of interesting motivational datasets, access to appropriate web-based tools for data manipulations and limited technical instructional skills are among the factors inhibiting the adoption of technology and integration of real datasets in the modern science education curricula (Dinov and Christou 2011).

This article presents a modern IT-blended interactive teaching approach for multivariate distributions, marginal, joint and conditional probabilities. We developed a pair of a learning activity and a dynamic webapp for the special case of bivariate Normal distribution, which demonstrate applied and theoretical multivariate concepts. Instructors may augment the applications presented here by using other continuous or discrete multivariate data (e.g. http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data). For instance, one can discuss the need for variable transformations, data distribution parametric assumptions and joint or conditional probabilities in the case of exploring monetary inflation and housing prices (using a 50-year SOCR Inflation vs. Housing Price Index dataset, http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_MonetaryBaseStocksInterest1959_2009). Open-source and community-based development, validation and support for computational resources, instructional materials and interactive learning activities are critical for training the next generation of skilled, productive and imaginative citizens.

Acknowledgments

We are indebted to all SOCR faculty, students and staff for their help, motivation and support for the development, deployment and validation of this bivariate distribution webapp. Motivation, ideas and source code from Kyle Siegrist (UAH), Dennis Pearl (OSU) and the Probability Distributome project (www.Distributome.org) were instrumental for these developments. This work was funded in part by NSF grants 1023115, 1022560, and 1022636, 0716055 and 0442992, and NIH grants P41EB015922 and UL1TR000124.

APPENDIX

In this appendix, we include some of the more theoretical considerations for defining, computing and applying marginal, conditional and bivariate normal probability calculations.

Random observations of continuous (e.g. time, longitude and latitude for a vehicle or a plane) or discrete processes (e.g. number of customers arriving at a store within a fixed timeframe, weight in kilograms and height in centimeters, as whole numbers, for a human subject, outcomes of a real or virtual bivariate die and coin of experiment) are commonly represented as stochastic measurements or variables. In a typical bivariate situation, the answer to an *a priori* scientific question may require computing the *joint probability density function* for the process, $f_{X,Y}(x,y)$. Let $A \subseteq \Omega_X$ and $B \subseteq \Omega_Y$ represent (non-trivial) subsets (e.g. intervals) of the ranges of the variables X and Y , Ω_X and Ω_Y , respectively. Assuming the joint density is known, the general form of the conditional probability of X given Y , i.e. $X|Y$, (for continuous variables) is given by

$$P(X \in A | Y \in B) = \frac{\int_{y \in B} \int_{x \in A} f_{X,Y}(x,y) dx dy}{\int_{y \in B} \int_{x \in \Omega_X} f_{X,Y}(x,y) dx dy}$$

The expression in the discrete variable case is similar with summation replacing the integration operators. In the special case where $B = \{y_0\}$, representing a single value, the conditional probability satisfies this equation

$P(X \in A | Y = y_0) = \frac{\int_{x \in \Omega_X} f_{X,Y}(x, y_0) dx}{\int_{x \in \Omega_X} f_{X,Y}(x, y_0) dx}$. Clearly, if the set (range) A is trivial, then the conditional probability is zero. A note of caution here is that if A is non-trivial and Y is a continuous variable, a problem arises when we attempt to define the conditional probability, $P(X|Y)$, as the probability $P(Y = y_0) = 0$, and both sides of the previous equation are trivial; this is the Borel–Kolmogorov paradox illustrating the ambiguity of defining conditional probability (Rao 1988). However, this problem may be circumvented by employing measure-theoretic definition of conditional expectation (Shafer and Vovk 2006; Billingsley 2012). Note that the joint density function could be defined more generally on any well-defined region in two-dimensional, not necessarily a Cartesian product of the marginal ranges, as we defined it earlier. The complexity of specifying the boundary (and the inside) of the support/domain of the bivariate density (using general Lebesgue integrals (Kiefer and Schwartz 1965; Stein 1981)) abstracts the notation and obfuscates the direct correspondence between the theoretical formulation and the computational application of the probability distribution and inference.

These are generic descriptions of the bivariate joint and conditional probabilities. Now, suppose that X has specifically Normal distribution, the conditional mean of X given $Y = y_0$, $E(X|Y = y_0)$, is linear in Y , and the conditional variance of X given $Y = y_0$, $\text{Var}(X|Y = y_0)$, is constant. Under these three predicaments, the conditional probability distribution of X given $Y = y_0$, $f_{X|Y=y_0}(X, Y) = f_{X|Y_0}(X, Y)$, is given by

$$f_{X|y_0} \sim N\left(\mu_{X|y_0} = \mu_X + \rho \frac{\sigma_X}{\sigma_Y} (y_0 - \mu_Y); \sigma_{X|y_0}^2 = \sigma_X^2 (1 - \rho^2)\right), \text{ (Gut 2009; Franke et al.}$$

2011) where $X \sim N(\mu_X; \sigma_X^2)$; the expectation and variance of Y are $E(Y) = \mu_Y$ and

$\text{Var}(Y) = E(Y^2) - \mu_Y^2 = \sigma_Y^2$; and the bivariate correlation

$$\rho = \text{Corr}(X, Y) = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y}. \text{ Note that } Y \text{ is not necessarily required itself to be}$$

normally distributed. However, if Y is also normally distributed, $Y \sim N(\mu_Y; \sigma_Y^2)$; we can explicitly compute the conditional and joint probabilities using the (marginal) density of Y ,

$$f_Y(y) = \frac{1}{\sigma_Y \sqrt{2\pi}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}. \text{ If we denote the conditional distribution of } X \text{ given } Y = y_o \text{ by}$$

$$\begin{aligned} g_{X|Y}(x|y_o) &= \frac{1}{\sigma_{X|Y} \sqrt{2\pi}} e^{-\frac{(x-\mu_{X|Y})^2}{2\sigma_{X|Y}^2}} \\ &= \frac{1}{\sigma_X \sqrt{1-\rho^2} \sqrt{2\pi}} e^{-\frac{\left(x-\mu_X - \rho \frac{\sigma_X}{\sigma_Y} (y_o - \mu_Y)\right)^2}{2\sigma_X^2 (1-\rho^2)}} \end{aligned}$$

Then, the *joint probability density function* of X and Y is *Bivariate Normal* and has the following closed form expression:

$$f_{X,Y}(x,y) = g_{X|Y}(x|y) f_Y(y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} e^{-q(x,y)}$$

where

$$q(x,y) = \frac{1}{2} \left(\frac{1}{1-\rho^2} \right) \left(\left(\frac{x-\mu_X}{\sigma_X} \right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X} \right) \left(\frac{y-\mu_Y}{\sigma_Y} \right) + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 \right)$$

The case of multivariate (three-dimensional or higher) joint, marginal and conditional probability distributions may be worked out similarly; however, the complexity added by using more than two variables may impact the *a priori* assumptions and make the calculations more intricate.

References

- Al-Aziz J, Christou N, et al. SOCR motion charts: an efficient, open-source, interactive and dynamic applet for visualizing longitudinal multivariate data. JSE. 2010; 18(3):1–29.
- Algozzine B, Bateman LR, et al. Developing technology competencies in a college of education. Current Issues in Education. 2011; 2(0):1–12.
- Ayalew L, Kasahara M, et al. The spatial correlation between earthquakes and landslides in Hokkaido (Japan), a GIS-based analysis of the past and the future. Landslides. 2011; 8:1–16.
- Billingsley, P. Probability and Measure. Wiley; 2012.
- Castillo E, Galambos J. Conditional distributions and the bivariate normal distribution. Metrika. 1989; 36(1):209–214.
- Che A, Cui J, et al. SOCR analyses: implementation and demonstration of a new graphical statistics educational toolkit. JSS. 2009; 30(3):1–19. [PubMed: 21666874]
- Christou N, Dinov I. Confidence interval based parameter estimation – a new SOCR applet and activity. PloS One. 2011; 6(5):e19178. [PubMed: 21655319]
- Dietz JR, Stevenson FT. Active learning in a large medical classroom setting for teaching renal physiology. Advances in Physiology Education. 2011; 35(4):456–459. [PubMed: 22139786]
- Dinov I. Statistics Online Computational Resource. Journal of Statistical Software. 2006; 16(1):1–16.

- Dinov I, Christou N. Statistics Online Computational Resource for Education. *Teaching Statistics*. 2009; 31(2):49–51. [PubMed: 21297884]
- Dinov I, Christou N. Web-based tools for modelling and analysis of multivariate data: California ozone pollution activity. *International Journal of Mathematical Education in Science and Technology (JMEST)*. 2011; 42(6):789–805.
- Franke, J.; Härdle, WK., et al. Basic concepts of probability theory. *Statistics of Financial Markets*. Springer; Berlin Heidelberg: 2011. p. 43–53.
- Franklin C, Kader G, et al. Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report. 2007
- Fraser D, Streit F. A further note on the bivariate normal distribution. *Communications in Statistics – Theory and Methods*. 1980; 9(10):1097–1099.
- Gudbjartsson DF, Walters GB, et al. Many sequence variants affecting diversity of adult human height. *Nature Genetics*. 2008; 40(5):609–615. [PubMed: 18391951]
- Gut, A. *An Intermediate Course in Probability*. Springer; 2009.
- Henneman EA, Gawlinski A, et al. Surveillance: a strategy for improving patient safety in acute and critical care units. *Critical Care Nurse*. 2012; 32(2):e9–e18. [PubMed: 22467622]
- Holbrook J, Rannikmae M. The nature of science education for enhancing scientific literacy. *International Journal of Science Education*. 2007; 29(11):1347–1362.
- Huang X, Dedegikas C, et al. Using multimedia technology to teach modern Greek language online in China: development, implementation, and evaluation. *European Journal of Open, Distance and E-Learning*. 2011; 9:1–9.
- Hwang G-J, Wu P-H, et al. An interactive concept map approach to supporting mobile learning activities for natural science courses. *Computers & Education*. 2011; 57(4):2272–2280.
- Jahnavi DG, Patra DSR. Customized teaching for better learning. *Journal of Nursing Science & Practice*. 2012; 2(1):1–11.
- Jordan J, Haines B. The role of statistics educators in the quantitative literacy movement. *JSE*. 2006; 14(2):1–12.
- Kiefer J, Schwartz R. Admissible Bayes character of T^2 , R^2 , and other fully invariant tests for classical multivariate normal problems. *Annals of Mathematical Statistics*. 1965; 36(3):747–770.
- Lancaster HO. Some properties of the bivariate normal distribution considered in the form of a contingency table. *Biometrika*. 1957; 44:289–292.
- Leung S, Lau J, et al. Weight-forage and weight-for-height references for Hong Kong children from birth to 18 years. *Journal of Paediatrics and Child Health*. 1996; 32(2):103–109. [PubMed: 8860382]
- Newton ER, Marshall PJ, et al. The Sloan Lens ACS Survey. XI. Beyond Hubble resolution: size, luminosity, and stellar mass of compact lensed galaxies at intermediate redshift. *The Astrophysical Journal*. 2011; 734:104.
- Patel, JK.; Read, CB. *Handbook of the Normal Distribution*. CRC; 1996.
- Rao M. Paradoxes in conditional probability. *Journal of Multivariate Analysis*. 1988; 27(2):434–446.
- Shafer G, Vovk V. The Sources of Kolmogorov's Grundbegriffe. *Statistical Science*. 2006; 21(1):70–98.
- Shelley, M.; Yore, L., et al. *Quality Research in Literacy and Science Education*. Springer; Netherlands: 2009. New directions in science literacy education.; p. 245–260.
- So HK, Nelson E, et al. Secular changes in height, weight and body mass index in Hong Kong Children. *BMC Public Health*. 2008; 8(1):320. [PubMed: 18803873]
- Stein CM. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*. 1981; 9:1135–1151.
- Stephens SH, Franks A, et al. Multiple genes in the 15q13-q14 chromosomal region are associated with schizophrenia. *Psychiatric Genetics*. 2012; 22(1):1. [PubMed: 21970977]
- Weedon MN, Lango H, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics*. 2008; 40(5):575–583. [PubMed: 18391952]

Fig. 1.

The Statistics Online Computational Resource (SOCR) bivariate Normal distribution web-application (webapp) user interface. The webapp is initiated using the settings parameters on the top-left. The control panel allows specification of the type and range of the probability density calculation. The graphs on the bottom illustrate the visual representations of the probabilities as areas under the curves/surfaces.

Fig. 2.
Distributions of adolescents' height (inches), left, and weight (pounds), right.

Mean = 67.94980003356933 variance = 3.7650069884913426	Mean = 127.22195022583008 variance = 143.0645467738852
Kolmogorov-Smirnoff Test for Differences Between the Data Distribution and the Model Distribution	Kolmogorov-Smirnoff Test for Differences Between the Data Distribution and the Model Distribution
Hypotheses: Ho: The Data follow the Model Distribution H1: The Data and Model Distributions are distinct	Hypotheses: Ho: The Data follow the Model Distribution H1: The Data and Model Distributions are distinct
KS D-Statistics = 0.033762 Z-Statistics = 0.238736 CDF(0.033762) = 0 P-value = 1	KS D-Statistics = 0.034653 Z-Statistics = 0.245037 CDF(0.034653) = 0.000001 P-value = 0.999999
Details about the Kolmogorov-Smirnoff Test are available here: http://webi.stat.ucla.edu/bsoc/mdeu.php?SOCR_EduMaterials_AnalysisActivities_KolmogorovSmirnoff	Details about the Kolmogorov-Smirnoff Test are available here: http://webi.stat.ucla.edu/bsoc/mdeu.php?SOCR_EduMaterials_AnalysisActivities_KolmogorovSmirnoff

Fig. 3.

Estimation of the mean and standard deviation for adolescents' height (left) and weight (right), and the corresponding Kolmogorov–Smirnov test quantitatively indicating the good fit between the data and the corresponding Normal distribution models.

Fig. 4. Using Statistics Online Computational Resource (SOCR) linear regression applet to compute the correlation $\rho = \text{Corr}(C2 = \text{height}, C3 = \text{weight})$.

Fig. 5.
Supports for the marginal distributions of X = height, left, and Y = weight, right.

Fig. 6.

Joint probability calculations: $P(65 < X < 70 \cap 120 < Y < 140) = 0.4894$.

Fig. 7.

Conditional probability calculations: $P(120 < Y < 140 | X = 67.95) = 0.661$.