implies that $(m, n+1) = 1$, so the above result becomes

$$F_{\Phi(m)+1} \equiv 1 (\operatorname{mod} m).$$

The other congruences follow similarly by letting $t = \Phi(m) - 1$ and $t = \Phi(m)$. ■

While the authors could not find the exact formulation of the theorem in the literature, we mention that the theorem can be shown to be equivalent to the following lemma of Koshy (see [**8**, pp. 408–409, Lemma 34.1]), which he proved using strong induction.

Lemma. *For integers $m \geq 2$ and $n \geq 0$,*

$$F_{n-1} - mF_n \equiv (-1)^n m^n (\operatorname{mod} m^2 + m - 1).$$

In conclusion, the reader is invited to extend the theorem to more general situations. Specifically, what congruence can be found concerning the solution to a more general recurrence like $G_n = cG_{n-1} + dG_{n-2}$, with arbitrary initial conditions G_0 and G_1, where c, d, G_0, and G_1 are natural numbers? Benjamin et al. tackled such questions using different phases for the initial tile and different colors for non-initial tiles. What if we allow arbitrary integers instead of just natural numbers? How about congruences concerning solutions of higher order difference equations? Can other ideas in the series [**1**]–[**3**] be carried over to find more congruences?

Acknowledgment. The authors would like to thank the anonymous referee for several suggestions that helped improve the quality of this exposition.

References

1. A.T. Benjamin, G.M. Levin, K. Mahlburg, and J.J. Quinn, Random approaches to Fibonacci identities, *Amer. Math. Monthly* **107** (2000) 511–516.
2. A.T. Benjamin and J.J. Quinn, The Fibonacci numbers—exposed more discretely, *Math. Mag.* **76** (2003) 182–192.
3. A.T. Benjamin, C.R.H. Hanusa, and F.E. Su, Linear recurrences through tilings and Markov chains, *Utilitas Math.* **64** (2003) 3–17.
4. A.T. Benjamin and J.J. Quinn, Recounting Fibonacci and Lucas identities, *College Math. J.* **30** (1999) 359–366.
5. A.T. Benjamin, J.J. Quinn, and F.E. Su, Phased tilings and generalized Fibonacci identities, *Fibonacci Quart.* **38** (2000) 282–288.
6. A.T. Benjamin and J.J. Quinn, *Proofs That Really Count: The Art of Combinatorial Proof*, MAA, 2003.
7. W.G. Kelley and A.C. Peterson, *Difference Equations: An Introduction with Applications*, 2nd ed., Academic Press, 2001.
8. T. Koshy, *Fibonacci and Lucas Numbers with Applications*, Wiley, 2001.

The Sample Correlation Coefficient from a Linear Algebra Perspective

C. Ray Rosentrater (rosentr@westmont.edu), Westmont College, Santa Barbara, CA 93013

In the undergraduate curriculum, the linear algebra course and the statistics course tend to have little to do with each other. This is unfortunate since a linear algebra perspective can provide a great deal of insight into statistical questions. In [**1**], Byer presents an analysis of the behavior of the sample correlation coefficient when one of the samples is nearly constant. He works with

$$r(\vec{x}, \vec{y}) = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\sum x^2 - \left(\sum x\right)^2}\sqrt{n\sum x^2 - \left(\sum y\right)^2}} \tag{1}$$

as the definition of the sample correlation coefficient, and he makes a series of observations based on the special vectors

$$\vec{x}(n) = (1, 2, \ldots, n) \quad \text{and} \quad \vec{y}_i(k, s) = (k, k, k, \ldots, k, k+s, k, \ldots, k)$$

where $k + s$ appears in position i.

Our goal is to re-examine Byer's observations from the vantage point of linear algebra. The lens of linear algebra will provide insight and structure not present in the original paper.

Let $\vec{x} = (x_1, x_2, \ldots, x_n)$ be a random sample. Depending on the context, we can view $\overline{x}$ as either a scalar or a scalar multiple of the vector $\vec{1} = (1, 1, \ldots, 1) \in \mathbb{R}^n$. In the later case, $\overline{x}$ is the projection of $\vec{x}$ into the one-dimensional space spanned by $\vec{1}$. This is verified implicitly in most statistics classes by showing that $\overline{x}$ is the value of c that minimizes

$$\sum_i (x_i - c)^2 = \left\|\vec{x} - c\vec{1}\right\|^2.$$

Alternatively, the result follows from a routine use of inner products.

Now define $\widehat{x} = \vec{x} - \overline{x}$. Note that we have decomposed the vector $\vec{x} = \widehat{x} + \overline{x}$ into two orthogonal components; the projection into the one-dimensional space spanned by $\vec{1}$ and its orthogonal complement. From this perspective, we immediately see that $\|\vec{x}\|^2 = \|\widehat{x}\|^2 + \|\overline{x}\|^2$. This provides insight into the equivalence of the two computational formulas for variance since

$$\frac{1}{n}\sum_i (x_i - \overline{x})^2 = \frac{1}{n}\|\widehat{x}\|^2 \quad \text{and} \quad \frac{1}{n}\sum_i x_i^2 - \overline{x}^2 = \frac{1}{n}(\|\vec{x}\|^2 - \|\overline{x}\|^2). \tag{2}$$

Continuing this projectional point of view, it is easy to see that if $\vec{p}$ is a perturbation of the data then $\overline{x + p} = \overline{x} + \overline{p}$ so that $\widehat{x + p} = \widehat{x} + \widehat{p}$ as well.

With $\langle u, v\rangle$ denoting the inner product, the sample correlation coefficient can be re-expressed as

$$r(\vec{x}, \vec{y}) = \frac{\langle\widehat{x}, \widehat{y}\rangle}{\|\widehat{x}\|\,\|\widehat{y}\|}, \tag{3}$$

which we note is the cosine of the angle between the vectors $\widehat{x}$ and $\widehat{y}$. It is immediately clear that $r(\vec{x}, \vec{y})$ is undefined when either $\vec{x}$ or $\vec{y}$ is a constant vector (that is, a multiple of $\vec{1}$) since, in those cases, $\widehat{x} = 0$ or $\widehat{y} = 0$. Byer's paper was motivated by the question of what happens when $\vec{y}$ is "almost" constant.

In the case where $\vec{y}$ is a perturbation of a constant vector, say $\vec{y} = k\vec{1} + \vec{p}$, we have $\widehat{y} = \widehat{p}$ and

$$r(\vec{x}, \vec{y}) = \frac{\langle \widehat{x}, \widehat{p} \rangle}{\|\widehat{x}\| \, \|\widehat{p}\|}. \tag{4}$$

Setting $\vec{p} = s\vec{u}$ where s is a positive scalar, (4) becomes

$$r(\vec{x}, \vec{y}) = \frac{\langle \widehat{x}, s\widehat{u} \rangle}{\|\widehat{x}\| \, \|s\widehat{u}\|} = \frac{\langle \widehat{x}, \widehat{u} \rangle}{\|\widehat{x}\| \, \|\widehat{u}\|}. \tag{5}$$

If s is a negative scalar, then the final expression is negated. It is worth noting that, since $\widehat{x}$ is orthogonal to the portion of $\vec{u}$ lying in the direction of $\vec{1}$, we have

$$r(\vec{x}, \vec{y}) = \frac{\langle \widehat{x}, \vec{u} \rangle}{\|\widehat{x}\| \, \|\widehat{u}\|}. \tag{6}$$

The following proposition of Byer is a special case of (5); it results from noting that $\vec{y}_i(k, s) = k\vec{1} + s\vec{e}_i$, where e_i is the ith standard basis element.

Proposition. Let $\vec{x} = (x_1, x_2, \dots, x_n)$ be a vector in $\mathbb{R}^n$. Then for $s \neq 0$,

$$r(\vec{x}, \vec{y}_i(k, s)) = -r(\vec{x}, \vec{y}_i(k, -s)),$$

and the values are independent of k and $|s|$.

Again, since $\widehat{x}$ is orthogonal to the component of $\vec{y}_i(k, s)$ in the direction of $\vec{1}$, we conclude from (6) that $r(\vec{x}, \vec{y}_i(k, s))$ will be a fixed multiple of the ith element of $\widehat{x}$. In the case of $\vec{x}(n) = (1, 2, \dots, n)$, the scalar factor is

$$\frac{1}{\|\widehat{x}(n)\| \, \|\widehat{e}_i\|} = \frac{1}{\sqrt{(n(n^2-1)/12)(1 - 1/n)}} = \frac{2}{(n-1)\sqrt{(n+1)/3}}.$$

Since

$$\overline{x}(n) = \frac{n+1}{2} \quad \text{and} \quad \widehat{x}(n) = \vec{x}(n) - \frac{n+1}{2}\vec{1},$$

the ith entry of $\widehat{x}(n)$ is

$$i - \frac{n+1}{2} = \frac{2i - n - 1}{2}.$$

This provides a more geometric explanation for the observation that

$$r(\vec{x}(n), \vec{y}_i(k, s)) = \frac{2i - n - 1}{(n-1)\sqrt{(n+1)/3}}.$$

When n is odd,

$$\widehat{x}(n) = \left(-\frac{n-1}{2}, -\frac{n-3}{2}, \dots, -1, 0, 1, \dots, \frac{n-1}{2}\right),$$

accounting for Byer's observation that

$$r(\vec{x}(n), \vec{y}_m(k, s)) = 0,$$

where m is the middle index.

Byer concludes by observing that for a fixed vector $\vec{x}$ and $\vec{y} = (y_1, y_2, \ldots, y_n)$,

$$\lim_{y_n \to \infty} r(\vec{x}, \vec{y}) = r(\vec{x}, \vec{y}_n(0, 1)).$$

Byer proves this fact by algebraically manipulating (1). Taking a linear algebra perspective, we can place this in a more general context and provide better intuition. Fix vectors $\vec{x}$, $\vec{y}$, and $\vec{u}$. Then use of (6) yields

$$\begin{aligned}
\lim_{s \to \infty} r(\vec{x}, \vec{y} + s\vec{u}) &= \lim_{s \to \infty} \frac{\langle \widehat{x}, \vec{y} + s\vec{u} \rangle}{\|\widehat{x}\| \, \|\widehat{y} + s\widehat{u}\|} \\
&= \lim_{s \to \infty} \frac{\langle \widehat{x}, \frac{1}{s}\vec{y} + \vec{u} \rangle}{\|\widehat{x}\| \, \|\frac{1}{s}\widehat{y} + \widehat{u}\|} \\
&= \frac{\langle \widehat{x}, \vec{u} \rangle}{\|\widehat{x}\| \, \|\widehat{u}\|} \\
&= r(\vec{x}, \vec{u}).
\end{aligned}$$

We have seen how ideas from linear algebra can give insight into the sample correlation coefficient. They can also provide intuition in the study of correlation and covariance of distributions. While none of these ideas are new or deep, they should enable the reader to appreciate the explanatory power of a linear algebra approach in a statistical context. Sadly, while some linear algebra texts like **[2]** include applications to statistics, available statistics texts, perhaps in an effort to reduce prerequisites, do not seem to relate concepts in statistics to ideas from linear algebra. Even **[3]**, which uses ideas from matrix theory in the context of analysis of variance and which proves and names a version of the Cauchy-Schwartz inequality in the context of the covariance of distributions, does not point out the relationship between the inner product and covariance. This is an unfortunate loss of a helpful tool for understanding covariance and correlation. It is always a good idea to find ways to reference, use, and reinforce ideas studied in other courses.

References

1. O. Byer, Observations on the Indeterminacy of the Sample Correlation Coefficient, *College Math. J.* **33** (2002) 316-318.
2. D. C. Lay, *Linear Algebra and its Applications*, 3rd ed., Addison-Wesley, 2003.
3. W. C. Rinaman, *Foundations of Probability and Statistics,*Saunders, 1993.

Pythagoras by the Cross Ratio

Rebecca M. Conley (rconley@austincollege.edu) and John H. Jaroma (jjaroma@austincollege.edu), Austin College, Sherman, TX 75090