Available online at www.sciencedirect.com

ScienceDirect

Journal of Statistical Planning and Inference 137 (2007) 945–952

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Extensions to Basu's theorem, factorizations, and infinite divisibility

Anirban DasGupta*

Purdue University, West Lafayette, IN, USA

Available online 1 August 2006

Abstract

We define a notion of approximate sufficiency and approximate ancillarity and show that such statistics are approximately independent pointwise under each value of the parameter. We do so without mentioning the somewhat nonintuitive concept of completeness, thus providing a more transparent version of Basu's theorem. Two total variation inequalities are given, which we call approximate Basu theorems.

We also show some new types of applications of Basu's theorem in the theory of probability. The applications are to showing that large classes of random variables are infinitely divisible (id), and that others admit a decomposition in the form YZ, where Y is infinitely divisible, Z is not, both are nondegenerate, and Y and Z are independent.

These applications indicate that the possible spectrum of applications of Basu's theorem is much broader than has been realized.

MSC: Primary 60E07; secondary 62F10

Keywords: Ancillary; Basu's theorem; Completeness; Factorization; Goldie–Steutel law; Independent; Infinitely divisible; Sufficiency; Total variation

1. Introduction

It has now been half a century since Basu (1955) proved what has turned out to be one of the most fundamental theorems of basic statistical theory. Popularly known as Basu's theorem, his result says that a boundedly complete and sufficient statistic is independent of an ancillary statistic under all values of the parameter. Although Basu's theorem is discussed in many texts, a particularly comprehensive and delightful review with many types of applications is given in Ghosh (2002). The purpose of this article is to obtain certain extensions to Basu's theorem and to offer a collection of new and probably even surprising applications of Basu's theorem.

The statistical intuition in Basu's theorem is that a statistic which captures all the information in a sample on an unknown parameter and another which captures none should provide no information about each other, and thus should be independent. The condition of completeness is in a way a necessary technical evil; one fails to see the intuition of requiring completeness; see, however, Lehmann (1981) and Ghosh (2002). It therefore seems natural to ask whether a sufficient statistic and an ancillary statistic would be 'nearly' independent anyway, in some well formulated sense. We ask, in fact, a slightly more general question: is it the case that an approximately sufficient statistic and an approximately

* Tel.: +1 765 494 6033; fax: +1 765 494 0558.
E-mail address: dasgupta@stat.purdue.edu.

0378-3758/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2006.06.018

ancillary statistic are approximately independent under all values of the parameter? Of course, approximate sufficiency and ancillarity would have to be defined. In Section 2, we give two total variation inequalities that imply such a result. We like to call such results *approximate Basu theorems*. Viewed in this manner, we think, the real intuition of Basu's theorem comes through in a more satisfactory manner because we can establish the 'near independence' without the somewhat abstract condition of completeness.

It is well known that although Basu's theorem is a theorem in statistical inference, it can be used to find easy solutions of distributional questions in probability. Ghosh (2002) gives many examples of applications of Basu's theorem in deriving joint distributions of random variables bypassing heavy calculations that would be needed in a direct attack. However, we give some entirely new types of applications. We show that a wide variety of random variables are infinitely divisible (id) by simultaneous use of Basu's theorem and an extended version of the Goldie–Steutel law of infinite divisibility (Goldie, 1967; Bose et al., 2002). The results do not follow from just the extended Goldie–Steutel law; Basu's theorem is also used. To our knowledge, Basu's theorem has never been applied prior to this in proving infinite divisibility. These results are presented in Section 3. The wide collection of explicit id densities obtained can be used for statistical modeling and as examples in classroom instruction.

In Section 4, we present another new type of application of Basu's theorem. We apply Basu's theorem in factorization of random variables, a topic of some interest in probability theory. The typical result we show says that a certain random variable X can be factorized as $X = YZ$, where Y is id, but Z is not, and Y and Z are independent.

In summary, the intention of this article is to give what we believe is a more transparent and more intuitive interpretation of Basu's theorem by avoiding the mention of completeness, and to demonstrate that the possible horizon of applications of Basu's theorem is probably much wider than has been understood so far.

It is a profound regret that this paper was not finished before Dev Basu passed away in the March of 2001. I offer these results as my homage to a remarkable scholar who imparted an ineffable influence on my life.

2. Approximate Basu theorems

The extensions to Basu's theorem are presented in this section. Each is given in terms of the total variation distance between a true joint distribution and the distribution that is the product of the corresponding marginals. For example, for specificity, suppose T is a sufficient statistic, and U an ancillary and θ is an unknown parameter. Let P_θ denote the joint distribution of T and U and Q_θ the product measure. Since T and U are independent under each θ if T is also boundedly complete, in such a case the total variation distance between P_θ and Q_θ would be zero.We ask what kinds of bounds can we establish on the total variation distance between P_θ and Q_θ without requiring completeness of T. For some of the results, we ask the same question, only taking T to be 'approximately sufficient' and U to be 'approximately ancillary'. Choice of total variation was a conscious choice as it seems very natural, but of course similar results should be possible with other distances, such as Kullback–Leibler or Hellinger.

First, we introduce some notation to be used for the rest of the article.

With respect to some common dominating measure $\mu \otimes v$, let $p_\theta(t, u) =$ joint density of (T, U); $f_\theta(u|t) =$ conditional density of U given $T = t$; $g_\theta(t) =$ marginal density of T; $h_\theta(u) =$ marginal density of U; $P_\theta =$ true joint distribution of (T, U); $Q_\theta =$ product measure corresponding to the marginals of T and U; $d_\theta =$ total variation distance between P_θ and Q_θ.

The exact choice of T and U will be given in the specific context.

Note that if T is sufficient and U is ancillary, then for a.a. $(t, u)(\mu \otimes v)$, $f_\theta(u|t)$ is independent of θ; i.e., there is a function $f(u, t)$ which acts as the conditional density under each θ. For the sake of notational simplicity, we will use the Lebesgue measure for each of μ and v.

Theorem 1. *Let T, U be general statistics. Then*

$$d_\theta \leqslant \int \sqrt{Var_\theta(f_\theta(u|T))}\, \mathrm{d}u.$$

Proof. By definition,

$$2d_\theta = \int |p_\theta(t, u) - g_\theta(t)h_\theta(u)|\, \mathrm{d}t\, \mathrm{d}u = \int |f_\theta(u|t) - h_\theta(u)|g_\theta(t)\, \mathrm{d}t\, \mathrm{d}u;$$

now observe that $E_{T|\theta}(f_\theta(u|T)) = h_\theta(u)$, and hence, $\int |f_\theta(u|t) - h_\theta(u)| g_\theta(t) \leqslant \sqrt{Var_\theta(f_\theta(u|T))}$; integrating this inequality over u gives the inequality of the theorem. □

Remark. If T is complete and sufficient and U is ancillary, then $Var_\theta(f_\theta(u|T)) = 0$ for almost all u, forcing $d_\theta = 0$, which is Basu's theorem. In this sense, the result above is an extension to Basu's theorem.

We illustrate Theorem 1 by an example.

Example 1. Suppose $X_1, X_2, \ldots, X_n$ are iid $U[0, \theta]$. Let $T = X_{(n)}$ and $U = X_{(1)}$. Note that T is sufficient (and even complete), but U is not ancillary. But intuitively, U has almost no information about θ, and so almost an ancillary. One might expect that T and U are almost independent. Let us see how Theorem 1 works out in this example.

Since T is sufficient, the conditional density of U given T is free of θ; denoting it by $f(u|t)$, by a direct calculation,

$$f(u|t) = (n-1)(t-u)^{n-2}/t^{n-1} I_{t>u}.$$

On the other hand, the density of T under θ of course is $g_\theta(t) = nt^{n-1}/\theta^n I_{0<t<\theta}$. Then on a few lines of algebra,

$$E_\theta(f(u|T)) = n(\theta-u)^{n-1}/\theta^n.$$

Similarly, $E_\theta(f(u|T)^2) = n(n-1)^2/\theta^n[(-1)^n \binom{2n-4}{n-2} u^{n-2}(\log\theta - \log u) + \sum_{j\neq n-2}(-1)^j \binom{2n-4}{j} u^j(\theta^{n-2-j} - u^{n-2-j})/(n-2-j)]$.

These expressions provide $Var_\theta(f(u|T))$. The bound of Theorem 1 is obtained by integrating the square root over $u \in (0, \theta)$. This integral can be very easily done numerically, but not in a closed form. We provide below a few illustrative values; the bound is for $\theta = 1$.

n	Bound of Theorem 1 on d_θ
5	.174
10	.079
15	.051
20	.040

The bounds are consistent with the intuition that $X_{(n)}$ and $X_{(1)}$ should be nearly independent for large n.

Next we give a result that shows that an approximately sufficient statistic and an approximately ancillary statistic are approximately independent. There is no mention of completeness in this result. First we specify a notion of approximate sufficiency and approximate ancillarity. We should mention that classic notions of approximate sufficiency have been around for a long time; see, e.g., Le Cam (1964, 1974, 1986), Reiss (1978), and Brown and Low (1996). It would be interesting if these notions of approximate sufficiency can be tied to the sort of approximate Basu theorems we want to establish. We have not attempted it.

Definition 1. A statistic T is called δ-sufficient with respect to another statistic U if $|\partial f_\theta(u|t)/\partial\theta| \leqslant \delta$ for a.a. (u, t) under each θ.

Definition 2. A statistic U is called ε-ancillary if $|\partial h_\theta(u)/\partial\theta| \leqslant \varepsilon$ for a.a. u under each θ.

Remark. Obviously, it is a part of the definitions that the stated partial derivatives exist. Clearly, sufficiency implies δ-sufficiency for any $\delta > 0$ and for any choice of U. Thus δ-sufficiency is a notion of approximate sufficiency, but only locally, in the sense it is with respect to the specified statistic U. Analogously, ε ancillarity with $\varepsilon = 0$ would mean ancillarity. Thus, the notions of approximate sufficiency and ancillarity given above are weaker than their usual meanings. However, we will see in the next theorem that one can obtain results in the spirit of Basu's theorem with these weaker notions.

One more remark is in order. Note that δ-sufficiency is invariant with respect to a one-to-one transformation on the statistic T. It is not invariant with respect to a monotone reparametrization. If we permit δ to depend on θ, then approximate sufficiency holds with a new value of δ. However, ultimately, the bound given in Theorem 2 below remains unchanged due to a cancellation with the Fisher information term under the reparametrization.

We now provide the theorem and its proof.

Theorem 2. *Let U be ε-ancillary and T δ-sufficient with respect to U. Suppose U and θ are bounded, taking values in, say, [0,1]. Assume $f_\theta(u|t)$, $g_\theta(t)$ and $h_\theta(u)$ are once continuously differentiable in θ for a.a. (u,t) and that $\lim_{\theta\to 0}[f_\theta(u|t)-h_\theta(u)]\sqrt{g_\theta(t)}=0$ for a.a. (u,t). Assume also that $\exists$ constants k_1, k_2 such that $f_\theta(u|t)\leqslant k_1$, and $h_\theta(u)\leqslant k_2$. Then,*

$$d_\theta \leqslant 1/\sqrt{2}\sqrt{2(\delta^2+\varepsilon^2)+(\max(k_1,k_2))^2\int I_g(\theta)\,\mathrm{d}\theta/4},$$

where $I_g(\theta)$ denotes the Fisher information in T about θ.

Remark. The interpretation of Theorem 2 is that if T is approximately sufficient and U approximately ancillary, then provided that T and U are independent under a 'degenerate' value of θ, they are approximately independent under all θ. There is no mention of completeness here. The conditions that U and θ belong to [0,1] can be relaxed to the conditions that they belong to bounded intervals.

Careful examination of the assumptions made in Theorem 2 reveals that collectively they imply that under some special value $(\theta=0)$, U and T are independent, and away from the special value, the relevant densities are not allowed to vary excessively. So a result of the sort of Theorem 2 would be expected. The interest in this kind of a result lies in a formulation of approximate sufficiency and approximate ancillarity that precipitate an approximate Basu theorem.

Proof. By definition,

$$\begin{aligned} 2d_\theta &= \int_0^1\int |f_\theta(u|t)-h_\theta(u)|g_\theta(t)\,\mathrm{d}t\,\mathrm{d}u \\ &\leqslant \sqrt{\int_0^1\left(\int |f_\theta(u|t)-h_\theta(u)|g_\theta(t)\,\mathrm{d}t\right)^2\mathrm{d}u} \quad \text{(by Schwartz's inequality applied to } u) \\ &\leqslant \sqrt{\int_0^1\int (f_\theta(u|t)-h_\theta(u))^2 g_\theta(t)\,\mathrm{d}t\,\mathrm{d}u} \quad \text{(by Schwartz's inequality applied to } T). \end{aligned}$$

Consider now the function $u=u(\theta)=(f_\theta(u|t)-h_\theta(u))\sqrt{g_\theta(t)}$. By the fundamental theorem and Schwartz,

$$u(\theta)\leqslant\sqrt{\int_0^1(\partial/\partial\theta\, u(\theta))^2\,\mathrm{d}\theta};$$

but,

$$\begin{aligned} \partial/\partial\theta\, u(\theta) &= (\partial/\partial\theta f_\theta - \partial/\partial\theta\, h_\theta)\sqrt{g_\theta} + (f_\theta-h_\theta)\partial/\partial\theta\, g_\theta/(2\sqrt{g_\theta}) \\ &\Rightarrow (\partial/\partial\theta\, u(\theta))^2 \leqslant 2[2(\delta^2+\varepsilon^2)g_\theta + [\max(k_1,k_2)]^2(\partial/\partial\theta\, g_\theta)^2/(4g_\theta)]; \end{aligned}$$

(by the simple Holder inequality $(a+b)^2\leqslant 2(a^2+b^2)$ and by the hypotheses of δ-sufficiency and ε-ancillarity).

Thus, $u^2(\theta)\leqslant 2\int_0^1[2(\delta^2+\varepsilon^2)g_\theta+(\max(k_1,k_2))^2(\partial/\partial\theta\, g_\theta)^2/(4g_\theta)]\,\mathrm{d}\theta$; note the important point that this integral bound on $u^2(\theta)$ is a *uniform* bound—it does not depend on θ. □

Now, since we have already proved that $2d_\theta\leqslant\sqrt{\int_0^1\int u^2\,\mathrm{d}t\,\mathrm{d}u}$, the theorem follows on integrating the above integral bound we derived on u^2 with respect to t, u and on performing the integrations in the order u, t, θ. The algebra is omitted.

3. Basu's theorem and infinite divisibility

In this section and the next, we show various applications of Basu's theorem to infinite divisibility and factorization of random variables. Some of the examples are known; but others are new. More than the new examples, the interesting thing is that Basu's theorem is useful in these kinds of problems. It should be pointed out that sampling distributional results needed for our results can be derived by textbook methods involving Jacobians. However, use of Basu's theorem leads to these results in a quicker and (we think) more elegant way. Instances of such quick consequences of Basu's theorem are eloquently documented in Boos and Hughes-Oliver (1998) and Ghosh (2002).

The results and the applications of the results depend on two facts. See Steutel (1970, 1979), and Bose et al. (2002). First we state these two facts as lemmas.

Lemma 1. *Let* $V \sim \mathrm{Exp}(1)$ *and W independent of V. Then the product VW is id.*

Remark. If W is nonnegative, then Lemma 1 is the Goldie–Steutel law which says that scale mixtures of Exponential densities are id. Essentially the same proof handles the case of a general W as well.

Lemma 2. *Let* $V \sim \mathrm{Exp}(1)$ *and suppose* $\alpha > 0$. *Then* V^α *admits the representation* $V^\alpha = UW$, *where* $U \sim \mathrm{Exp}(1)$, *and W is independent of U.*

We now state and prove our main theorems of this section.

Theorem 3. *Let f be any homogeneous function of two variables, i.e., suppose* $f(cx, cy) = c^2 f(x, y)\ \forall x, y$ *and* $\forall c > 0$. *Let* Z_1, Z_2 *be iid* $\mathrm{N}(0, 1)$ *random variables and* $Z_3, Z_4, \ldots, Z_m$ *any other random variables such that* $(Z_3, Z_4, \ldots, Z_m)$ *is independent of* (Z_1, Z_2). *Then for any positive integer k, and an arbitrary measurable function g,* $f^k(Z_1, Z_2)g(Z_3, Z_4, \ldots, Z_m)$ *is id.*

Due to the fact that f can be any homogeneous function and k and g are completely arbitrary, in principle, Theorem 3 has a very broad range of applications. Before giving a proof of Theorem 3, we state a corollary of this theorem and show a fairly large number of examples as applications of Theorem 3. The examples establish a large variety of random variables as being id.

Corollary 1. (a) *Let* $f(x, y)$ *be any of the functions* xy, $x^2 + y^2$, $|x|^\alpha |y|^\beta$ *where* $\alpha, \beta \geqslant 0$ *and* $\alpha + \beta = 2$, $\sqrt{x^4 + y^4}$, $(x^n + y^n)/(x^{n-2} + y^{n-2})$ *where* $n \geqslant 2$.

Then for iid $\mathrm{N}(0, 1)$ *random variables* Z_1, Z_2, *and* $Z_3, Z_4, \ldots, Z_m$ *any other random variables such that* $(Z_3, Z_4, \ldots, Z_m)$ *is independent of* (Z_1, Z_2), *any positive integer k, and an arbitrary measurable function g,* $f^k(Z_1, Z_2)$ $g(Z_3, Z_4, \ldots, Z_m)$ *is id.*

(b) *Let* $n \geqslant 1$ *and* $X_1, X_2, \ldots, X_n$ *independent random variables, each a* scale mixture *of zero mean normal distributions. Then for any* $k \geqslant 1$, $(X_1 X_2 \ldots X_n)^k$ *is id.*

(c) *Let* $m, n \geqslant 1$, $m < n$ *and* $X_1, X_2, \ldots, X_n$ *independent random variables, each a* scale mixture *of zero mean normal distributions. Then for any* $p, q \geqslant 1$, *the ratio* $R = (X_1 X_2 \ldots X_m)^p / (X_{m+1} \ldots X_n)^q$ *is id.*

Proof of Corollary 1. We will first prove part (a); parts (b) and (c) essentially follow from part (a).

Observe that each function f mentioned in part (a) is a homogeneous function as is easily verified. Thus, directly from Theorem 3, part (a) follows as a corollary.

To see the result of part (b), write $X_1 X_2 \ldots X_n = (\sigma_1 Z_1)(\sigma_2 Z_2) \ldots (\sigma_n Z_n)$, where the $\{Z_i\}$ are $\mathrm{N}(0, 1)$ random variables and the $\{Z_i\}$ and the $\{\sigma_i\}$ are all mutually independent. Therefore, $X_1 X_2 \ldots X_n = Z_1 Z_2 (Z_3 \ldots Z_n \sigma_1 \sigma_2 \ldots \sigma_n)$, and so it follows from part (a) that $X_1 X_2 \ldots X_n$ is id by taking f to be xy, k as 1, and g as the product function in the corresponding space. The proof for a general k is the same.

Part (c) follows exactly similarly and so we omit the proof. □

We will now give, as specific illustrative examples, density functions that correspond to id distributions. It should be remarked that some are known, and others new. Densities (i), (iii), (v) and (viii) in the example below are simple functions and can be used as models in applications or as examples in classroom instruction.

Example 2. Each of the following density functions on $(-\infty, \infty)$ correspond to id distributions.

(i) $f(x) = 1/\pi K_0(|x|)$, where K_0 denotes the Bessel K_0 function;
(ii) $f(x) = -\exp(x^2/2)Ei(-x^2/2)/(\sqrt{2}\pi^{3/2})$, where $Ei(.)$ denotes the Exponential integral;
(iii) $f(x) = K_0(2\sqrt{|x|})$;
(iv) $f(x) = [\sin(|x|)(\pi - 2\mathrm{si}(|x|)) - 2\cos(x)\mathrm{ci}(|x|)]/(2\pi)$, where si(.) and ci(.) stand for the sine and the cosine integral;
(v) $f(x) = 2\log(|x|)/(\pi^2(x^2-1))$;
(vi) $f(x) = [|x| - \sqrt{2\pi}\exp(1/(2x^2))\Phi(-1/|x|)]/(\sqrt{2\pi}|x|^3)$, where $\Phi(.)$ denotes the N(0, 1) CDF;
(vii) $f(x) = (1 - \sqrt{2\pi}|x|\exp(x^2/2)\Phi(-|x|))/\sqrt{2\pi}$;
(viii) $f(x) = 1/(2(1+|x|)^2)$;
(ix) $f(x) = \exp(1/(2x^2))\Gamma(0, 1/(2x^2))/(\sqrt{2}\pi^{3/2})$,where $\Gamma(.,.)$ denotes the incomplete Gamma function;
(x) $f(x) = [\sin(1/|x|)(\pi - 2\mathrm{si}(1/|x|)) - 2\cos(1/|x|)\mathrm{ci}(1/|x|)]/(2\pi x^2)$.

Proof. It follows as a special case of parts (b) and (c) of Corollary 1 that the densities stated in Example 2 are all id. For the purpose of this example, we will denote a standard normal, a standard double exponential, and a standard Cauchy random variable as N, D, and C, respectively. Then the densities (i)–(x) are, respectively, the density functions of NN, NC, DD, CD, CC, N/D, D/N, D/D, C/N and C/D, where the notation NN means the product of two independent standard normals, etc. Therefore, the infinite divisibility of each one directly follows from parts (b) and (c) of Corollary 1. □

Remark. Note that the density of N/C is the same as that of NC, and the density of D/C is the same as that of CD. Thus, they are not separately mentioned in the example. And of course, N/N is the same as C, and therefore not mentioned either.

Proof of Theorem 3. Let Z_1, Z_2 be iid N(0, 1) and $(Z_3, \ldots, Z_m)$ independent of (Z_1, Z_2). Consider any homogeneous function $f(Z_1, Z_2)$ and write it as $f(Z_1, Z_2) = (Z_1^2 + Z_2^2)f(Z_1, Z_2)/(Z_1^2 + Z_2^2)$. Basu's theorem will now be used. Introduce, just for the sake of the proof, a parameter σ^2 and take the more general case where Z_1, Z_2 are iid $\mathrm{N}(0, \sigma^2)$. Then $Z_1^2 + Z_2^2$ is complete and sufficient and $f(Z_1, Z_2)/(Z_1^2 + Z_2^2)$ is ancillary because f is homogeneous. Therefore, $f(Z_1, Z_2)/(Z_1^2 + Z_2^2)$ and $Z_1^2 + Z_2^2$ are independent under *any* σ, and so in particular, under $\sigma = 1$. Thus, $f(Z_1, Z_2)$ can be written as VW, where $V \sim \mathrm{Exp}(2)$ and W is independent of V.

Hence, $f^k(Z_1, Z_2)g(Z_3, \ldots, Z_m) = V^k W^k g(Z_3, \ldots, Z_m)$. Now apply Lemma 2. By Lemma 2, $f^k(Z_1, Z_2)g(Z_3, \ldots, Z_m) = V^k W^k g(Z_3, \ldots, Z_m) = UW_* W^k g(Z_3, \ldots, Z_m)$, where $U \sim \mathrm{Exp}(1)$ and the rest are independent of U, and therefore by the extended Goldie–Steutel Law (Lemma 1), $f^k(Z_1, Z_2)g(Z_3, \ldots, Z_m)$ is id.

4. Factorization of random variables

In this final section, we will show that Basu's theorem can be used to decompose various functions of iid N(0, 1) random variables in the form *YZ* where *Y* is id, *Z* is not, both are nondegenerate, and *Y* and *Z* are independent. Although they are all functions of iid N(0, 1) random variables, the class of functions that admit such a decomposition is large, as we see in the following theorem.

Theorem 4. *Let $X_1, X_2, \ldots, X_n$ be iid* N(0, 1) *random variables. Suppose $h_i(x_1, x_2, \ldots, x_n)$, $1 \leqslant i \leqslant n$, are scale invariant functions, i.e., $h_i(cx_1, cx_2, \ldots, cx_n) = h_i(x_1, x_2, \ldots, x_n)$ $\forall c > 0$, and f is any continuous homogeneous function in the n-space, i.e., $f(cx_1, cx_2, \ldots, cx_n) = c^n f(x_1, x_2, \ldots, x_n)$ $\forall c > 0$.*

Define $W = g(X_1, X_2, \ldots, X_n) = f(\|X\|h_1(X_1, X_2, \ldots, X_n), \ldots, \|X\|h_n(X_1, X_2, \ldots, X_n))$, where $\|\mathbf{X}\|$ denotes the Euclidean norm of the vector $(X_1, X_2, \ldots, X_n)$.

Then W admits the representation $W = YZ$, where Y is id, Z is not, both are nondegenerate, and Y and Z are independent.

Before giving a proof of Theorem 4, we give a few interesting examples of such a decomposition that would follow from the theorem.

Example 3. Suppose $X_1, X_2, \ldots, X_n$ are iid N(0, 1); then each of the following random variables can be decomposed as *YZ*, with Y and Z as in Theorem 4:

(i) $W = X_1 X_2 \ldots X_n$;
(ii) $W = X_1^n + X_2^n + \cdots + X_n^n$;
(iii) $W = X_1^2 X_2^2 \ldots X_n^2/(X_1^2 + X_2^2 + \cdots + X_n^2)^{n/2}$;
(iv) $W = (X_1^{2n} + X_2^{2n} + \cdots + X_n^{2n})/(X_1^2 + X_2^2 + \cdots + X_n^2)^{n/2}$.

Of these four specific examples, it is particularly nice that (i) and (ii) have the decomposition we are discussing because they have a particular special form. That each of these four random variables has the stated decomposition can be seen by proving two more general facts. Let us do that now.

Consider the scale invariant functions $h_i(x_1, x_2, \ldots, x_n) = x_i^m/(x_1^2 + x_2^2 + \cdots + x_n^2)^{m/2}$, and the continuous homogeneous function $f(x_1, x_2, \ldots, x_n) = x_1^n + x_2^n + \cdots + x_n^n$.

Then,

$$\begin{aligned} f(\|x\|h_1(x_1, x_2, \ldots, x_n), \ldots, \|x\|h_n(x_1, x_2, \ldots, x_n)) &= \|x\|^n \sum_{i=1}^{n} h_i^n(x_1, x_2, \ldots, x_n) \\ &= \|x\|^n \sum_{i=1}^{n} x_i^{mn}/\|x\|^{mn} \\ &= \sum_{i=1}^{n} x_i^{mn}/\|x\|^{n(m-1)}. \end{aligned}$$

The special value $m = 1$ gives the random variable in (ii), and the special value $m = 2$ gives the random variable in (iv).

For the other two examples, consider the same scale invariant functions h_i, but change the function *f* to $f(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n$.

Then, $f(\|x\|h_1(x_1, x_2, \ldots, x_n), \ldots, \|x\|h_n(x_1, x_2, \ldots, x_n)) = \|x\|^n (x_1 x_2 \ldots x_n)^m/(x_1^2 + x_2^2 + \cdots + x_n^2)^{(nm/2)}$.

The special value $m = 1$ gives the random variable in (i), while $m = 2$ gives the random variable in (iii).

Of course, numerous other examples can be worked out by simply choosing other functions h_i and *f*.

We will finish by giving a proof of Theorem 4.

Proof of Theorem 4. By definition of g, and from the homogeneity of *f*, we have that $g(X_1, X_2, \ldots, X_n) = \|X\|^n f(h_1(X_1, X_2, \ldots, X_n), \ldots, h_n(X_1, X_2, \ldots, X_n))$.

Of these, $\|X\|^n$ is a power of a chi-square and hence id (this is well known). As regards the second factor, if we introduce as before an artificial parameter σ^2 and let $X_1, X_2, \ldots, X_n$ be iid N(0, σ^2), then $\|X\|$ is complete and sufficient, and the vector of functions $U = (h_1(X_1, X_2, \ldots, X_n), \ldots, h_n(X_1, X_2, \ldots, X_n))$ is ancillary because by hypothesis each h_i is scale invariant. *Hence*, by Basu's theorem $\|X\|$ and U are independent, under any σ, and so in particular under $\sigma = 1$. So if we let $Y = \|X\|^n$, and $Z = f(U)$, then *Y* and *Z* are independent. To see that *Z* cannot be id, note that by its scale invariance, *f* is determined by its values on the unit disk and therefore must be bounded as *f* was also assumed to be continuous. Since a bounded random variable cannot be id, it now follows that $g(X_1, X_2, \ldots, X_n)$ has the decomposition stated in the theorem. □

Acknowledgment

I am thankful to two anonymous referees, Stefan Ralescu, Gutti Jogesh Babu, and Herman Rubin for their detailed reading and very concrete suggestions. Both referees corrected a number of misstatements and algebra errors. They have my sincere appreciation.

References

Basu, D., 1955. On statistics independent of a complete sufficient statistic. Sankhyā 15, 377–380.

Boos, D., Hughes-Oliver, J.M., 1998. Applications of Basu's theorem. Amer. Statist. 52 (3), 218–221.

Bose, A., DasGupta, A., Rubin, H., 2002. A contemporary review and bibliography of infinitely divisible distributions and processes, Invited Article for the Basu Memorial Volume. Sankhyā, Sankhyā Ser. A 64 (3, Part II), 763–819.

Brown, L., Low, M., 1996. Asymptotic equivalence of nonparametric regression and white noise. Ann. Statist . 24 (6), 2384–2398.

Ghosh, M., 2002. Basu's Theorem with applications: a personalistic review, Invited Article for the Basu Memorial Volume. Sankhyā Ser. A 64 (3, Part I), 509–531.

Goldie, C., 1967. A class of infinitely divisible random variables. Proc. Cambridge Philos. Soc. 63, 1141–1143.

Le Cam, L., 1964. Sufficiency and approximate sufficiency. Ann. Math. Statist. 35, 1419–1455.

Le Cam, L., 1974. Construction of asymptotically sufficient estimates in some nonGaussian situations. Proc. Prague Sympos. Asymp. Statist. 1, 179–200.

Le Cam, L., 1986. Asymptotic Methods in Statistical Decision Theory. Springer, New York.

Lehmann, E.L., 1981. An interpretation of completeness and Basu's theorem. J. Amer. Statist. Assoc. 76, 335–340.

Reiss, R., 1978. Approximate Distributions of Order Statistics. Springer, New York.

Steutel, F.W., 1970. Preservation of infinite divisibility under mixing and related topics. Math Centre Tracts, vol. 33, Mathematisch Centrum, Amsterdam.

Steutel, F.W., 1979. Infinite divisibility in theory and practice. Scand. J. Statist. 6 (2), 57–64.