Elisabetta Bona; Dario Sacchetti
The inverse distribution for a dichotomous random variable

Commentationes Mathematicae Universitatis Carolinae, Vol. 38 (1997), No. 2, 385--394

Comment.Math.Univ.Carolin. 38,2 (1997)385–394

The inverse distribution for a dichotomous random variable

Elisabetta Bona, Dario Sacchetti

Abstract. In this paper we will deal with the determination of the inverse of a dichotomous probability distribution. In particular it will be shown that a dichotomous distribution admit inverse if and only if it corresponds to a random variable assuming values $(0, a),\ a \in \mathbb{R}^+$.

Moreover we will provide two general results about the behaviour of the inverse distribution relative to the power and to a linear transformation of a measure.

Keywords: inverse measure, inverse probability distribution, Laplace transform, variance function

Classification: 62E10

1. Introduction

Let X be a random variable and $P(x)$ its probability distribution; if the moment generating function (MGF) exists, then let denote it by $M_X(t)$ and the cumulant generating function (CGF) by $K_X(t) = \log M_X(t)$.

If

$$T = \{t \in \mathbb{R} : K'(t) > 0\} \neq \emptyset,$$

then $K(t)$ is a one-to-one function on T. In this case let

$$\widetilde{K}(t) = -K^{(-1)}(-t),$$
$$\widetilde{M}(t) = e^{\tilde{K}(t)}$$

and $\widetilde{P}(x)$ the inverse Laplace transform of $\widetilde{M}(t)$. Note that with $f^{(-1)}$ we will indicate the inverse function of f, i.e. $f \circ f^{(-1)} =$ *identity function.*

If $\widetilde{P}(x)$ is a probability distribution *i.e.* is always positive with total mass equal to one, then $\widetilde{P}(x)$ is called the inverse distribution of $P(x)$.

Alternatively, if X is a discrete random variable taking values on $\{1, 2, \dots\}$, let $G_X(t) = M_X(\log t)$ denote the probability generating function; in this case we have

$$\widetilde{G}(t) = \widetilde{M}[\log t] = \left[\frac{1}{G(1/t)}\right]^{(-1)}.$$

If $\widetilde{G}(t)$ can be expanded as a power series around the origin, that is

$$\widetilde{G}(t) = \sum_h a_h t^h$$

with $a_h \geq 0$, then

$$\widetilde{P}(x) = \sum_h a_h \delta_h(x)$$

and $\widetilde{P}(x)$ is the inverse distribution of $P(x)$. The definition of inverse distribution was introduced by Tweedie [9]; the Gaussian and the Inverse Gaussian are the most popular examples of inverse distributions. Moreover the binomial distribution with parameters (p, N) has as inverse the distribution of X/N where X is distributed as a geometric with parameter p; the inverse of the gamma distribution with parameters (p, N) is the distribution of X/N where X is a Poisson with parameter p ([7]).

Let observe that, given a probability distribution, the existence of its inverse is not in general guaranteed, since $\widetilde{P}(x)$ is not always (strictly) positive. For instance, if $P(x)$ is the distribution of the logarithmic series (LSD),

$$P(x,\theta) = \sum_{n=1}^{\infty} \frac{\theta^n}{n \log(1-\theta)} \delta_n(x),$$

where $\theta \in (0,1)$ and $\delta_n(x)$ is the Dirac function in n, we have

$$\widetilde{P}(x,\theta) = \theta \Big\{ \frac{1}{|\log(1-\theta)|} \delta_1(x) + \frac{1}{2}\delta_0(x) + \\ + \sum_{n=1}^{\infty} (-1)^{n-1} \frac{|B_{2n}|}{(2n)!} |\log(1-\theta)|^{2n-1} \delta_{-(2n-1)}(x) \Big\},$$

where B_{2n} are the Bernoulli numbers. In this case $\widetilde{P}(x,\theta)$ is a signed measure whose total mass is still equal to one ([6]).

The application that defines $\widetilde{P}(x)$ keeps some of the original properties of $P(x)$. The following results hold:

Proposition 1.1.

(i) *Both $P(x)$ and $\widetilde{P}(x)$ have total mass equal to one;*
(ii) *both $K(t)$ and $\widetilde{K}(t)$ are convex functions;*
(iii) *both $M(t)$ and $\widetilde{M}(t)$ are analytical functions.*

PROOF: (i) $\widetilde{M}(t) = \int_{\mathbb{R}} e^{tx}\, d\widetilde{P}(x)$ therefore $\widetilde{M}(0) = \int_{\mathbb{R}} d\widetilde{P}(x)$, but

$$\widetilde{M}(0) = e^{M^{①}}(e^{-t})_{t=0} = 1.$$

(ii) $K(t) = \log M(t)$ where $M(t)$ denotes the Laplace transform of $P(x)$. Both $M(t)$ and $K(t)$ are infinitely times differentiable. Convexity of $K(t)$ follows directly from Hölder's inequality; by the $\widetilde{K}(t)$ definition it is

$$\widetilde{K}'(t) = \frac{1}{K'(K^{(-1)}(-t))}$$

and

$$\widetilde{K}''(t) = \frac{K''(K^{(-1)}(-t))}{[K'(K^{(-1)}(-t))]^3}\,.$$

Since $K''(t) > 0$ and $K'(t) > 0$, it is $\widetilde{K}''(t) > 0$ too.

(iii) As it is known, $M(t)$ is analytic in his natural domain; since $\widetilde{M}(t)$ is obtained by composing analytic functions, it results that $\widetilde{M}(t)$ is an analytic function in a certain neighbourhood of the origin. □

Remark. The initial assumption $K' > 0$ is justified since, as shown in (ii), it implies the convexity of $\widetilde{K}$, a necessary condition for $\widetilde{K}$ to be the CGF of a positive measure.

With respect to the probability distributions in the natural exponential family (NEF), the problem of existence of the inverse distribution can be approached in an alternative way.

Let μ be a positive, not necessarily bounded measure, and assume that:

- $M_\mu(\theta) = \int_{\mathbb{R}} e^{\theta x}\, d\mu(x)$ is the Laplace's transform of μ;
- $K_\mu(\theta) = \log M_\mu(\theta)$;
- $D_\mu = \{\theta : M_\mu(\theta) < +\infty\}$ is the domain of $M_\mu(\theta)$;
- Θ_μ is the interior of D_μ.

Also, let us suppose that μ is not degenerate and that $\Theta_\mu \neq \emptyset$.

For each $\theta \in \Theta_\mu$, let

$$P_\mu(\theta)dx = \frac{e^{\theta x}}{M_\mu(\theta)}\mu(dx)$$

and let

$$F = \{P_\mu(\theta), \theta \in \Theta\}$$

be the NEF generated by μ; μ is called *basis* of F_μ.

It is well known that $K_\mu(\theta)$ is convex and, therefore, that $K'_\mu(\theta)$ is invertible.

Let E_F be the image of Θ_μ through the application $K'_\mu(\theta)$ and let $\psi(m)$ the inverse function of $K'_\mu(\theta)$; E_μ is called *mean domain* of F_μ.

Definition 1.1. *The application $V_\mu : E_\mu \to \mathbb{R}$ defined as*

$$V_\mu(m) = \int_{\mathbb{R}} (x-m)^2 P_\mu(m)\, dx, \quad m \in E_\mu$$

is called variance function of the NEF F.

The importance of the variance function relies on the fact that any NEF is uniquely determined by its domain and by the variance function itself ([5]). The following results are well known ([3]).

Theorem 1.1. *Let F_μ be the NEF generated by μ, and V_μ its variance function. Then:*

(i) $V_\mu(m) > 0 \ \ \forall\, m \in E_\mu$;
(ii) $V_\mu(m) = K''_\mu(\psi(m)) = \frac{1}{\psi'(m)} \ \ \forall\, m \in E_\mu$;
(iii) $V_\mu(m)$ *is analytical in* E_μ.

Theorem 1.2. *Consider*

$$\varphi(x) = ax + b, \quad a \neq 0,\ b \in \mathbb{R}$$

and let F_μ be the NEF generated by μ, μ_1 the image measure of μ through φ, and F_{μ_1} the NEF generated by μ_1; then:

(i) $E_{\mu_1} = \varphi(E_\mu)$;
(ii) $V_{\mu_1}(m) = a^2 V_\mu[(\frac{m-b}{a})] \quad \forall\, m \in E_{\mu_1}$.

Analogously to the probability distribution case, it is possible to introduce the concept of *inverse measure.*

Definition 1.2. *Let μ and $\tilde{\mu}$ be two non degenerate measures such that Θ_μ and $\Theta_{\tilde{\mu}}$ are nonempty; let us define*

$$\Theta_\mu^+ = \left\{\theta \in \Theta_\mu : K'_\mu(\theta) > 0\right\}.$$

The measure $\tilde{\mu}$ is the inverse measure of μ if:

(i) Θ_μ^+ *is nonempty;*
(ii) $-K_{\tilde{\mu}}(-K_\mu(\theta)) = \theta \quad \forall\, \theta \in \Theta_\mu^* \subset \Theta_\mu^+$, *where* $\Theta_\mu^* \neq \emptyset$.

If μ and $\tilde{\mu}$ are positive measures then the following theorem provides the connection between the variance functions V_μ and $V_{\tilde{\mu}}$.

Theorem 1.3. *Let F_μ and $\widetilde{F}_\mu$ be the NEF's generated by the inverse measures μ and $\tilde{\mu}$ respectively. Then*

(i) *the application* $m \to 1/m$ *is one-to-one from* $E_\mu \cap (0, +\infty)$ *to* $E_{\tilde{\mu}} \cap (0, +\infty)$;
(ii) $V_\mu(m) = m^3 V_{\tilde{\mu}}(1/m)$.

The problem of inversion of a NEF has been considered by several authors, classifying the NEF's according to their variance functions, computing the variance function of the inverse and then identifying the probability distribution.

Such a classification is exhaustive for the NEF whose variance function is a polynomial of degree less or equal to three [3], [4], [5].

2. Dichotomous measure

In this section we will consider the problem of the inversion of a dichotomous measure using two different methods: directly, i.e. using the inverse Laplace transform, or using the variance function.

We will determine the conditions for which the inverse of a dichotomus measure is positive and it will be then calculated, in an explicit manner, the inverse probability distribution.

First consider the following theorem

Theorem 2.1 (Lagrange's formula). *Let g be analytic in $(-r, r)$, $r > 0$ and $g(0) \neq 0$, then there exist an $R > 0$ and an analytic function $t = t(y)$ on $(-R, R)$ such that $t = y\, g(t)\ \forall\, y \in (-R, R)$.*

Furthermore, if F analytic on $(-r, r)$, then for all $y \in (-R, R)$ it is:

$$F(t) = F(0) + \sum_{n=1}^{\infty} \frac{y^n}{n!} \left\{ D^{n-1} \left[F'(t)(g(t))^n \right] \right\}_{t=0}.$$

Proof: See for example [1]. □

Theorem 2.2. *Let $\mu = \delta_a + \delta_b$, with $a, b \in \mathbb{R}^+$, then the inverse measure $\tilde{\mu}$ exists if and only if $a = 0$.*

Proof: 1° *method:*

If $\mu = \delta_a + \delta_b$, $a, b \in \mathbb{R}^+$, $a < b$, then

$$G(t) = t^a + t^b, \quad t > 0 \quad \text{and} \quad \frac{1}{G(1/t)} = \frac{t^b}{1 + t^{b-a}}.$$

Let $w = [G(1/t)]^{-1}$ then $t^b = w\left(1 + t^{b-a}\right)$. Let $z = t^{b-a}$ then $z^{\frac{b}{b-a}} = w(1+z)$ e.g. $z = w^{\frac{b-a}{b}}(1+z)^{\frac{b-a}{b}}$. The function $g(z) = (1+z)^{\frac{b-a}{b}}$ is analytic in $(-1, 1)$ and $g(0) = 1$ then for Theorem 2.1, when F is the identity function and $y = w^{\frac{b-a}{b}}$, it is:

$$z = \sum_{n=1}^{\infty} \frac{w^{n\frac{b-a}{b}}}{n!} \left\{ D^{n-1} (g(z))^n \right\}_{z=0}.$$

Since

$$(g(z))^n = (1+z)^{n\frac{b-a}{b}} = \sum_{h=0}^{+\infty} \binom{n\frac{b-a}{b}}{h} z^h,$$

it is

$$D^{n-1}(g(z))^n_{z=0} = \binom{n\frac{b-a}{b}}{n-1}(n-1)!$$

Then

$$z=\sum_{n=1}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}(n-1)!\frac{w^{n\frac{b-a}{b}}}{n!}=\sum_{n=1}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}\frac{w^{n\frac{b-a}{b}}}{n}$$
$$=w^{\frac{b-a}{b}}\left\{1+\sum_{n=2}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}\frac{w^{(n-1)\frac{b-a}{b}}}{n}\right\}\qquad \forall\, w\in(-R;R).$$

Note that

$$\binom{n\frac{b-a}{b}}{n-1}=\frac{1}{(n-1)!}n\frac{b-a}{b}\left(n\frac{b-a}{b}-1\right)\dots\left(n\frac{b-a}{b}-n+2\right)$$

and

$$\binom{n\frac{b-a}{b}}{n-1}<0 \;\text{ for }\; a\neq 0,\; n=\left\lfloor\frac{2b}{a}\right\rfloor+1,$$
$$\binom{n\frac{b-a}{b}}{n-1}>0 \qquad \forall\, n\in\mathbb{N}\iff a=0,$$

where $\lfloor x\rfloor$ is the integer part of x. If $a>0$, then the quantity

$$(*)\qquad t=z^{\frac{1}{b-a}}=w^{1/b}\left\{1+\sum_{n=2}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}\frac{w^{(n-1)\frac{b-a}{b}}}{n}\right\}^{\frac{1}{b-a}}$$
$$=w^{1/b}\sum_{h=0}^{+\infty}\binom{\frac{1}{b-a}}{h}\left[\sum_{n=2}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}\frac{w^{(n-1)\frac{b-a}{b}}}{n}\right]^{h}$$

will be of the type

$$w^{1/b}\left\{\sum_{r=0}^{+\infty}a_r w^{\frac{b-a}{b}r}\right\}\quad\text{for}\quad w\in\left(-R';R'\right).$$

Let observe that if $a_r>0$, $\forall\, r\in\mathbb{N}$, the coefficients of the serie $\left\{\sum_{r=0}^{+\infty}a_r w^{\frac{b-a}{b}r}\right\}^{b-a}$ will be all positive, but, as already seen, the coefficients of the serie

$$\left\{\sum_{r=0}^{+\infty}a_r w^{\frac{b-a}{b}r}\right\}^{b-a}=\left\{1+\sum_{n=2}^{+\infty}\binom{n\frac{b-a}{b}}{n-1}\frac{w^{n\frac{b-a}{b}}}{n}\right\}$$

are all positive only if $a=0$.

Finally, since

$$G(w)=w^{1/b}\left\{\sum_{r=0}^{+\infty}a_r w^{\frac{b-a}{b}r}\right\}$$

for Proposition (45.2) in [2], it is

$$\tilde{\mu} = \delta_{\frac{1}{b}} \left\{ \sum_{r=0}^{+\infty} a_r \delta_{\frac{b-a}{b} r} \right\} = \sum_{r=0}^{+\infty} a_r \delta_{\frac{1+(b-a)}{b} r}$$

and $\tilde{\mu}$ comes out to be a signed measure.

If $a = 0$ the $(*)$ becomes

$$t = z^{1/b} = w^{1/b} \left\{ 1 + \sum_{n=2}^{+\infty} \binom{n}{n-1} \frac{w^{(n-1)}}{n} \right\}^{1/b}$$
$$= w^{1/b} \left[\sum_{n=1}^{+\infty} w^{(n-1)} \right]^{1/b} = w^{1/b} (1-w)^{-1/b}.$$

Therefore

$$\widetilde{G}(w) = w^{1/b} \left\{ \sum_{h=0}^{+\infty} \binom{-1/b}{h} w^h \right\}$$

and

$$\tilde{\mu} = \sum_{h=0}^{+\infty} \binom{-1/b}{h} \delta_{\frac{1}{b}+h}$$

is, as already seen, a positive measure.

2° *method:*

Let $\mu_0 = \delta_0 + \delta_1$, F_{μ_0} the NEF associated to μ_0 and V_{μ_0} the variance function of F_{μ_0}. Then ([3]),

$$E_{\mu_0} = (0,1) \quad \text{and} \quad V_{\mu_0} = m - m^2.$$

Now, let $\mu = \delta_a + \delta_b$ with $a, b \in \mathbb{R}^+ - \{0\}$, $a < b$, and F_μ the NEF associated to μ. Then $\mu = \varphi(\mu_0)$, where φ is the affine transform

$$\varphi(x) = (b-a)x + a.$$

From Theorem 1.2, we have

$$E_\mu = \varphi((0,1)) = (a,b) \quad \text{and} \quad V_\mu = (m-a)(b-m).$$

Now, if $F_{\tilde{\mu}}$ is the NEF associated to $\tilde{\mu}$, from Theorem 1.3 it follows that

$$E_{\tilde{\mu}} = (\frac{1}{b}, \frac{1}{a}).$$

Furthermore

$$V_{\tilde{\mu}} = m(1-ma)(mb-1).$$

However, the Mora-Morris classification [3] assures that no exponential family exists with variance function equal to a third degree polynomial and bounded domain. Therefore, μ does not have an inverse distribution.

Let us now consider the other cases.

Suppose $a = 0$ and $b > 0$. Then:

$$E_{\tilde{\mu}} = \left(\frac{1}{b}, \infty\right) \quad \text{and} \quad V_{\tilde{\mu}}(m) = m(bm - 1)$$

that is the variance function of the inverse binomial distribution with parameter $p = 1/b$ up to the affine tranformation $\varphi(x) = x + 1/b$ ([3]).

Finally, let $a = 0$ and $b < 0$, then:

$$E_{\tilde{\mu}} = (-\infty, \frac{1}{b}) \quad \text{and} \quad V_{\tilde{\mu}} = m(bm - 1);$$

this is absurd, since $V_{\tilde{\mu}}$ turns out to be negative. Therefore, the inverse measure $\tilde{\mu}$ of a measure μ exists if and only if $a = 0$. □

Corollary 2.1. *Let X be a random variable that assumes the values 0 and $a > 0$ respectively with probability p and $1 - p$. The inverse random variable, $\widetilde{X}$, is discrete and assumes the values $\{1/a + n - 1,\ n \in \mathbb{N}\}$ with probabilities respectively:*

$$(-1)^{n-1}\binom{-1/a}{n-1}p^{n-1}(1-p)^{1/a}.$$

Proof: Since it is

$$G(t) = p + (1-p)t^a,$$

$$\widetilde{G}(t) = (1-p)^{1/a}t^{1/a}(1-pt)^{-1/a}, \qquad |t| < \frac{1}{p},$$

then

$$\widetilde{P}(x) = (1-p)^{1/a}\sum_{n=1}^{+\infty}(-1)^{n-1}\binom{-1/a}{n-1}p^{n-1}\delta_{\frac{1}{a}+n-1}$$

and the corollary is proved. □

3. Inverse measures transformations

In this section we will describe the general behaviour of the inverse distributions with respect to the power and the linear transformation of a measure μ.

Theorem 3.1. *The following results hold:*

(i) $\widetilde{M_\mu^{1/a}}(t) = \widetilde{M}_\mu(at)$,

(ii) $\widetilde{G_\mu^{1/a}}(t) = \widetilde{G}_\mu\left(t^a\right)$.

Proof: It is, omitting for the sake of simplicity the indication of μ:

$$\widetilde{M}(t) = e^{M^{①}}\left(e^{-t}\right),$$

then

$$\widetilde{M^{1/a}}(t) = e^{-\left(M^{1/a}\right)^{①}\left(e^{-t}\right)} = e^{-M^{①}}\left(e^{-t}\right)^{a} = \widetilde{M}(at)$$

and (i) is proved. Since $G(t) = M(\log t)$, $t \in \mathbb{R}^{+}$, then (ii) follows. □

Theorem 3.2. *The following results hold:*

(i) $\widetilde{M}_{a\mu}(t) = \widetilde{M}_{\mu}^{1/a}(t)$,

(ii) $\widetilde{G}_{a\mu}(t) = \widetilde{G}_{\mu}^{1/a}(t)$.

Proof:

$$\widetilde{M}_{a\mu}(t) = e^{-M^{①}}\left(e^{-t}\right) = e^{-\frac{M_{\mu}^{①}\left(e^{-t}\right)}{a}} = \widetilde{M}_{\mu}^{1/a}(t),$$

then (i) is proved and (ii) follows. □

Observation 3.1. If X is a random variable such that $P\{X = 0\} = p$ and $P\{X = a\} = 1 - p$, $a \in \mathbb{N}$, for the Corollary 2.1, $\widetilde{X}$ assumes non integer values for $a \neq 1$.

Let $a \in \mathbb{N}$, $a \neq 1$, and $\widetilde{G}_{\widetilde{X}}(t) = [(1-p)t]^{1/a}(1-pt)^{-1/a}$, for (ii) of Theorem 3.1 it results:

$$\widetilde{G^{1/a}}(t) = \widetilde{G}\left(t^{a}\right) = (1-p)^{1/a}t(1-pt)^{-1/a} = $$
$$= (1-p)^{1/a}t\sum_{h=0}^{\infty}(-1)^{h}\binom{-1/a}{h}p^{h}t^{ah}$$

that is the generating function of the random variable assuming the integer values $1 + ah$, $h \in \mathbb{N}$, and having the same distribution of $\widetilde{X}$. Let observe that $G^{1/a}(t)$ is the $\frac{1}{a}$-*th* power measure of the probability distribution of the random variable X, according to the definition reported in [8, p. 37].

References

[1] Dieudonné J., *Infinitesimal Calculus*, Houghton Hifflin, Boston, 1971.

[2] Guest P.B., *Laplace transforms and an introduction to distributions*, Series mathematics and its applications, Ellis Horwood, 1991.

[3] Mora M., *Classification des functions-variance cubiques des families exponentielles*, C.R. Acad. Sci. Paris 302 sér. 1, **16** (1986), 582–591.

[4] Letac G., Mora M., *Natural real exponential families with cubic variance functions*, The Annals of Statistics **18** (1990), 1–37.

[5] Morris C.N., *Natural exponential families with quadratic variance functions*, The Annals of Statistics **10** (1982), 65–80.

[6] Sacchetti D., *Inverse distributions: an example of non existence*, Accademia di scienze, lettere ed arti di Palermo, 1993.

[7] Serrecchia A., *Inverse distributions*, Publ. de l'Inst. de Stat. de l'Univ. de PARIS XXXI **1** (1986), 71–85.
[8] Seshadri V., *The Inverse Gaussian Distribution*, Clarendon Press, Oxford, 1993.
[9] Tweedie H.C.K., *Inverse statistical deviates*, Nature (1945), 155–453.

Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università di Roma "La Sapienza", Roma, Italy

(*Received* May 3, 1996)