AN ABSTRACT OF THE THESIS OF

MICHAEL SPENCER WATERMAN for the M.S. in Mathematics
(Name) (Degree) (Major)

Date thesis is presented May 12, 1966

Title THE EXPONENTIAL FAMILY OF PROBABILITY DISTRIBUTIONS GENERATED BY σ-FINITE MEASURES

Abstract approved [redacted]
(Major professor)

The exponential family of probability distributions is obtained from σ-finite measures on the real line. We choose the parameter space to be one-dimensional and the exponent to be linear. Relationships between the measure, its spectrum, and the parameter space are examined, moments of the exponential family are studied, and the nature of a mapping of the parameter space into the spectrum is discussed. Examples of the many distributions belonging to the exponential family are given, including the binomial, Poisson, and normal with known variance. Finally, atomic probability distributions are considered, and a characterization of a certain subclass of these distributions is given.

THE EXPONENTIAL FAMILY OF PROBABILITY DISTRIBUTIONS GENERATED BY σ-FINITE MEASURES

by

MICHAEL SPENCER WATERMAN

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1966

APPROVED:

[redacted]

Associate Professor of Mathematics and Statistics

In Charge of Major

[redacted]

Chairman of Department of Mathematics

[redacted]

Dean of Graduate School

Date thesis is presented May 12, 1966

Typed by Carol Baker

ACKNOWLEDGEMENT

The author is indebted to Professor Donald Guthrie for his assistance during the preparation of this paper.

TABLE OF CONTENTS

THE EXPONENTIAL FAMILY OF PROBABILITY DISTRIBUTIONS GENERATED BY σ-FINITE MEASURES

INTRODUCTION

The exponential family of probability distributions can be obtained from σ-finite measures on the real line. This family of distributions is of interest in mathematical statistics and is usually studied in that context; however, the exponential family is considered here from the standpoint of elementary probability and measure theory.

In general, given a σ-finite measure μ, we will consider probability measures P_ω satisfying

$$P_\omega(A) = \beta(\omega) \int_A e^{\omega x} d\mu(x) \qquad (1)$$

where $\beta(\omega)$ is a normalizing factor. The density of P_ω, with respect to μ, is

$$p_\omega(x) = \beta(\omega) e^{\omega x} \quad . \qquad (2)$$

To see the statistical application of the exponential family, let $X_1, X_2, \cdots, X_n$ be a sample from a distribution with density (2). Then $\sum_{i=1}^{n} X_i$ is a minimal sufficient statistic for ω.

Conversely, let $X_1, X_2, \cdots, X_n$ be a sample from a distribution with a density p_ω, and suppose that $\{x: p_\omega(x) > 0\}$ is independent of ω. Then, according to Lehmann [6], under certain regularity conditions, if there exists a one-dimensional sufficient statistic, then the densities p_ω compose an exponential family. Also, as an application of likelihood ratio, there exists a UMP test for testing $H_1: \omega \leq \omega_0$ against $H_2: \omega > \omega_0$.

We define the spectrum of the measure μ, $\Lambda(\mu)$, to be the smallest (closed) interval that contains all the points x such that $\mu(U) > 0$ for all open sets U containing x. The parameter space, $\Omega(\mu)$, is the set of points ω such that (1) has meaning. (These definitions are made precise below.)

First some relationships between μ, $\Lambda(\mu)$, and $\Omega(\mu)$ are examined. Whenever $\Lambda(\mu)$ is bounded on the left, $\Omega(\mu)$ is unbounded on the left. Thus whenever $\Lambda(\mu)$ is a bounded set, $\Omega(\mu) = R$. It is known that $\Omega(\mu)$ is convex [6]. Finally, conditions equivalent to $\Omega(\mu)$ being bounded are given.

Next it is shown that all moments of P_ω exist if ω belongs to the interior of $\Omega(\mu)$. Denoting the mean of p_ω by $m(\omega)$, $m(\omega)$ is a continuous increasing function on the interior of $\Omega(\mu)$, and its range is contained in $\Lambda(\mu)$. Surprisingly enough the range of $m(\omega)$ frequently contains the interior of $\Lambda(\mu)$, and the main result of this thesis, Theorem 1.10, states when this occurs.

Many well known distributions belong to the exponential family and some of these are given here, including the binomial, Poisson, and negative binomial distributions on the integers and the exponential, Gamma, and normal with known variance in the absolutely continuous case.

Finally, Chapter 2 concerns discrete or atomic probability distributions. A characterization of exponential family distributions on the non-negative integers with positive weights at 0 is given. A very interesting subclass of these distributions found in [4] is presented and discussed.

For our discussion, we will have a fixed measurable space (R, Σ) where R denotes the real line and Σ the σ-algebra of Borel subsets of R; when measures on I, the integers, or N, the natural numbers, are considered, one could reduce the σ-algebra. However, the use of (R, Σ) is sufficient for all practical purposes, and the use of a fixed measurable space eliminates needless changes of the underlying space.

The set theoretic notation and definition of terms from measure theory will be found in [5] and are, for the most part, omitted here.

CHAPTER 1. THE EXPONENTIAL FAMILY

1.1 Definition of the Exponential Family

Consider an arbitrary σ-finite measure μ on R (so that (R, Σ, μ) is now a measure space). We make the added assumption that B is a bounded subset of $R, B \in \Sigma$, such that

$$\mu(B) > 0. \tag{1.1}$$

This restricts μ to nontrivial measures. Next define a and b by

$$\begin{aligned} a &= \inf\{x : \mu([x, x+\varepsilon)) > 0 \quad \text{for all} \quad \varepsilon > 0\}, \\ b &= \sup\{x : \mu((x-\varepsilon, x]) > 0 \quad \text{for all} \quad \varepsilon > 0\}. \end{aligned} \tag{1.2}$$

Here $a = -\infty$ or $b = +\infty$ is allowed. If a and b are finite, define the spectrum of the measure $\mu, \Lambda(\mu)$, by

$$\Lambda(\mu) = [a, b]. \tag{1.3}$$

If a or b are infinite, use open or half open intervals in equation (1.3).

The parameter space of μ, $\Omega(\mu)$, is defined by

$$\Omega(\mu) = \{\omega : 0 < \int e^{\omega x} d\mu(x) < +\infty\}. \tag{1.4}$$

This is the major step in defining the exponential family. The parameter space has been considered as a subset of the extended real line [1] but such an approach admits parameters which do not, in general, specify probability distributions. We have chosen the parameter space to be one-dimensional and the exponent to be linear in x. For a more general definition of the exponential family see [6]. By the existence of the set B in (1.1) we have

$$\int e^{\omega x} d\mu(x) \geq \int_B e^{\omega x} d\mu(x) \geq \inf_{x \in B} \{e^{\omega x}\} \mu(B) > 0$$

for all $\omega \epsilon R$, so that finiteness of the integral in (1.4) is the primary concern. For convenience we write, for all $\omega \epsilon \Omega(\mu)$,

$$\begin{aligned} n(\omega) &= \int e^{\omega x} d\mu(x), \\ \beta(\omega) &= [n(\omega)]^{-1}. \end{aligned} \tag{1.5}$$

The objective of our discussion is to generate probability measures from certain σ-finite measures. To accomplish this we define a distribution function, $F(\cdot\,;\omega)$, for each $\omega \epsilon \Omega(\mu)$ by

$$F(x;\omega) = \frac{\int_{(-\infty, x]} e^{\omega t} d\mu(t)}{\int e^{\omega t} d\mu(t)} . \tag{1.6}$$

The mean of this distribution (if it exists) is

$$m(\omega) = \frac{\int xe^{\omega x} d\mu(x)}{\int e^{\omega x} d\mu(x)} \quad . \tag{1.7}$$

For each $F(\cdot\,;\omega)$, there is a probability measure P_ω such that equation (1) holds. It is clear that we could write

$$m(\omega) = \int x dP_\omega(x).$$

The set of distribution functions and the set of probability measures are equivalent; we choose to define the exponential family to be the set of distribution functions satisfying (1.6) for some σ-finite measure μ.

1.2 Relationships between $\Lambda(\mu)$ and $\Omega(\mu)$

This section is a study of the relationships between the spectrum of a σ-finite measure μ and the parameter space $\Omega(\mu)$. Some of the results, besides being of independent interest, will have application to the problem discussed in 1.3.

Proposition 1.1. Suppose $\Omega(\mu)$ is non-empty. Then, if A is any bounded subset of R, $\mu(A) < +\infty$.

Proof: There exists $\omega \in \Omega(\mu)$. Clearly

$$0 < y = \inf \{e^{\omega x} : x \in A\} < +\infty$$

and

$$+\infty > \int e^{\omega x} d\mu(x) \geq \int_A e^{\omega x} d\mu(x) \geq y\ \mu(A).$$

Dividing by y yields $\mu(A) < +\infty$.

Next we consider certain conditions on $\Lambda(\mu)$ and make some conclusions about μ and $\Omega(\mu)$. The first consideration allows us to characterize a certain class of probability distributions.

Proposition 1.2. Assume a is finite and $\Lambda(\mu) = [a, a]$. Then $\Omega(\mu) = R$ and μ determines an improper distribution with jump at $x = a$.

Proof: Since μ is σ-finite, $0 < \mu(\{a\}) < +\infty$. We have

$$\int e^{\omega x} d\mu(x) = e^{\omega a} \mu(\{a\}) < +\infty , \quad \omega \in R,$$

which implies $\Omega(\mu) = R$.

$$F(x;\omega) = \beta(\omega) \int_{(-\infty, x]} e^{\omega t} d\mu(t) = \frac{1}{\mu(\{a\})} \int_{(-\infty, x]} e^{\omega(t-a)} d\mu(t)$$

$$= \begin{cases} 0, & x < a, \\ 1, & x \geq a . \end{cases} \qquad (1.8)$$

We have established the intuitively obvious fact that the set of measures satisfying the hypothesis of Proposition 1.2 determines all improper distributions on R.

Having fixed the measure μ, we then consider $F(\cdot\,;\omega)$ for $\omega \in \Omega(\mu)$. Denoting the variance of this distribution by $\sigma^2(\omega)$, we have the following proposition.

Proposition 1.3. Suppose there is some $\omega \in \Omega(\mu)$ such that $\sigma^2(\omega) = 0$. Then $\Lambda(\mu) = [m(\omega), m(\omega)]$.

Proof: $$\sigma^2(\omega) = \beta(\omega) \int (x - m(\omega))^2 e^{\omega x} d\mu(x) = 0.$$

Thus $x - m(\omega) = 0$ a.e. $[\mu]$, and $m(\omega)$ is the a of Proposition 1.2.

It follows from Proposition 1.3 that $\omega \in \Omega(\mu)$ such that $\sigma^2(\omega) = 0$ implies $\sigma^2(\omega) = 0$ for all $\omega \in \Omega(\mu)$. A similar result is given in [8].

Proposition 1.4. Assume $\Omega(\mu) \neq \phi$. If $b < +\infty$, then $\Omega(\mu)$ is an interval, unbounded on the right. If $a > -\infty$, then $\Omega(\mu)$ is an interval, unbounded on the left.

Proof: Assume $\omega_1 \in \Omega(\mu)$ and $a > -\infty$. Whenever $\omega_2 < \omega_1$ and $x \in [0, \infty)$, $e^{\omega_2 x} \leq e^{\omega_1 x}$. Thus we have

$$\int_{[0,\infty)} e^{\omega_2 x} d\mu(x) \leq \int_{[0,\infty)} e^{\omega_1 x} d\mu(x) \leq \int e^{\omega_1 x} d\mu(x) < +\infty .$$

If $a \geq 0$, then $\omega_2 \epsilon \Omega(\mu)$ and

$$(-\infty, \omega_1] \subset \Omega(\mu). \tag{1.9}$$

If $a < 0$, then, by Proposition 1.1, $\mu([a, 0)) < +\infty$. Since $e^{\omega_2 x}$ is a bounded function on $[a, 0)$, we have

$$\int e^{\omega_2 x} d\mu(x) < +\infty$$

which implies $\omega_2 \epsilon \Omega(\mu)$ and (1.9) holds.

The proof of the corresponding result for $b < +\infty$ is analagous.

Proposition 1.5. Assume $\Omega(\mu) \neq \phi$. If $\Lambda(\mu)$ is a bounded subset of R (that is, $-\infty < a$ and $b < +\infty$), then $\Omega(\mu) = R$.

Proposition 1.5 follows directly from either Proposition 1.1 or 1.4. Whenever $0 \epsilon \Omega(\mu)$, we have $\mu(R) < +\infty$. It should also be noted that, while $\Lambda(\mu)$ bounded implies $\mu(R) < +\infty$, $\Omega(\mu) = R$ does not follow from $\mu(R) < +\infty$. This is illustrated by the following example.

Example 1.1. Define a measure on N by

$$\mu(\{k\}) = k^{-n} \quad \text{for} \quad k \in N,$$

where $n \in N$. For $n = 1$ we have

$$\Omega(\mu) = \{\omega : 0 < \sum_{k=1}^{\infty} e^{\omega k} k^{-1} < +\infty\} = (-\infty, 0)$$

and $\mu(R) = \mu(N) = \sum_{k=1}^{\infty} \frac{1}{k} = +\infty$.

Notice that for $n \geq 2$ we obtain $\Omega(\mu) = (-\infty, 0]$ and $\mu(R) < +\infty$.

The example is also of interest because, for $n \geq 2$, $\Omega(\mu)$ is not an open set in R. If we consider the moments of the resulting distributions, it is clear that, for $n \geq 2$, there are exactly $n - 2$ moments corresponding to $\omega = 0$. However, all moments exist in the interior of $\Omega(\mu)$. This is a very suggestive property of this example and that it is true in general for the exponential family will be shown in 1.3.

The following theorem can be found in [6]. The proof is short and is also given.

Theorem 1.1. The parameter space of μ, $\Omega(\mu)$, is a convex set.

Proof: If $\Omega(\mu) = \phi$ or $\Omega(\mu) = \{\omega_0\}$, there is nothing to show.

Otherwise we have $\omega_1, \omega_2 \epsilon \Omega(\mu)$ such that $\omega_1 \neq \omega_2$. Take $0 < a < 1$ $(a + (1-a) = 1)$.

$$\int \exp\{(\omega_1 a + \omega_2(1-a)x\} d\mu(x) = \int e^{\omega_1 ax} e^{\omega_2(1-a)x} d\mu(x)$$

$$\leq \{ \int e^{\omega_1 x} d\mu(x)\}^{a} \{ \int e^{\omega_2 x} d\mu(x)\}^{1-a} \tag{1.10}$$

by Holder's inequality. Each term of (1. 10) is finite which implies $\omega_1 a + \omega_2(1-a) \epsilon \Omega(\mu)$. Therefore $\Omega(\mu)$ is a convex set.

An obvious restatement of the theorem is that $\Omega(\mu)$ is always an interval. Some authors have not realized this fact. Propositions 1. 2, 1. 4, and 1. 5 lead to the conclusion (independent of Theorem 1. 1) that $\Omega(\mu)$ is an interval having special properties due to the assumptions on $\Lambda(\mu)$.

Next, certain assumptions are made about $\Omega(\mu)$ and conclusions derived concerning μ and $\Lambda(\mu)$. The definition of $\Omega(\mu)$, equation (1. 4), leads us to a close consideration of $n(\omega)$, equation (1. 6).

We first inquire whether there are σ-finite measures μ such that $\Omega(\mu)$ is empty; above we have frequently assumed that this was not the case. The existence of such measures is shown by example.

<u>Example 1.2.</u> Let μ be a measure such that

$$\frac{d\mu}{dm}(x) = \begin{cases} e^{x^2}, & x > 0, \\ 0, & x \le 0 \end{cases} \tag{1.11}$$

where m denotes, as it shall throughout, Lebesgue measure on R. Consider, for any $\omega \in R$,

$$\int e^{\omega x} d\mu(x) = \int_{(0,\infty)} e^{x^2 + \omega x} dm(x).$$

This integral diverges so $\Omega(\mu) = \phi$.

We next consider measures that have bounded parameter spaces. Suppose we have a measure μ such that, for all $\varepsilon > 0$,

$$\int e^{\omega x} d\mu(x) < +\infty, \quad \omega_0 < \omega < \omega_1, \tag{1.12}$$

$$\int_{R^+} e^{(\omega_1 + \varepsilon)x} d\mu(x) = +\infty, \tag{1.13}$$

and

$$\int_{R^-} e^{(\omega_0 - \varepsilon)x} d\mu(x) = +\infty, \tag{1.14}$$

where $R^+ = \{x: x \ge 0\}$ and $R^- = R - R^+$. Define $M(\omega_0, \omega_1)$ to be the set of measures satisfying (1.12), (1.13), and (1.14). The following proposition is simply a restatement of definition and finds use in determining whether $\Omega(\mu)$ is bounded. We denote the interior of the set A by $\text{Int}(A)$.

<u>Proposition 1.6.</u> $\text{Int}\Omega(\mu) = (\omega_0, \omega_1)$ if and only if $\mu \in M(\omega_0, \omega_1)$.

<u>Proof:</u> (1.13) and (1.14) with the identity

$$\int f d\mu = \int_{R^+} f d\mu + \int_{R^-} f d\mu$$

imply $\text{Int}\Omega(\mu) \subset (\omega_0, \omega_1)$. (1.12) assures us that equality holds.

Conversely, suppose $\text{Int}\Omega(\mu) = (\omega_0, \omega_1) \neq \phi$. Since we have

$$(\omega_0 - \varepsilon)x \leq (\omega_0 + \varepsilon)x, \qquad x \in R^+,$$

and

$$(\omega_1 - \varepsilon)x > (\omega_1 + \varepsilon)x, \qquad x \in R^- ;$$

$$\int_{R^+} e^{(\omega_0 - \varepsilon)x} d\mu(x) \leq \int_{R^+} e^{(\omega_0 + \varepsilon)x} d\mu(x) \tag{1.15}$$

and

$$\int_{R^-} e^{(\omega_1 - \varepsilon)x} d\mu(x) \geq \int_{R^-} e^{(\omega_1 - \varepsilon)x} d\mu(x). \tag{1.16}$$

For $0 < \varepsilon < \omega_1 - \omega_0$, $(\omega_0 + \varepsilon) \in \Omega(\mu)$ and

$$\int_{R^+} e^{(\omega_0 - \varepsilon)x} d\mu(x) \leq \int e^{(\omega_0 + \varepsilon)x} d\mu(x) < +\infty .$$

Thus (1.15) yields (1.14). Similarly from (1.16) we have (1.13). (1.12) follows directly.

$M(\omega_0, \omega_1)$, $\omega_0 \leq \omega_1$, is non-empty as the following example shows.

Example 1.3. Define a measure μ on I by

$$\mu(\{k\}) = \begin{cases} 0, & k = 0, \\ (|k|^{1+\delta} e^{\omega_1 k})^{-1}, & k > 0, \\ (|k|^{1+\delta} e^{\omega_0 k})^{-1}, & k < 0, \end{cases} \qquad (1.17)$$

for $\delta \geq 0$. Let $J = I - \{0\}$.

Suppose $\omega_1 < \omega_1 + \varepsilon = \omega$. Then

$$\sum_{k \in J} \mu(\{k\}) e^{\omega k} \geq \sum_{k=1}^{\infty} \frac{e^{\varepsilon k}}{k^{1+\delta}} = +\infty .$$

Similarly

$$\sum_{k \in J} \mu(\{k\}) e^{\omega k} = +\infty \qquad \text{when} \quad \omega < \omega_0 .$$

But when $\omega_0 < \omega < \omega_1$,

$$0 < \sum_{k \in J} \mu(\{k\}) e^{\omega k} < +\infty .$$

The end points, ω_0 and ω_1, are in $\Omega(\mu)$ if $\delta > 0$. If $\omega_0 = \omega_1$, we require $\delta > 0$ so that $\Omega(\mu) \neq \phi$. The measure in this example has no moments at ω_0 or ω_1 if $\delta \leq 1$ and has exactly n moments if $n + 1 \geq \delta > n$. Obvious variations on this example allow ω_0 or ω_1 to belong to $\Omega(\mu)$.

For the case $\Omega(\mu) = \{\omega_0\}$, we have certain properties listed below.

<u>Corollary</u>. Suppose $\Omega(\mu) = \{\omega_0\}$. Then

(i) $\int_{R^+} e^{(\omega_0 - \varepsilon)x} d\mu(x) < +\infty$;

(ii) $\int_{R^-} e^{(\omega_0 + \varepsilon)x} d\mu(x) < +\infty$;

(iii) If $\omega_0 > 0$, then $0 < \mu(R^+) < +\infty$ and $\mu(R^-) = +\infty$;

(iv) If $\omega_0 < 0$, then $0 < \mu(R^-) < +\infty$ and $\mu(R^+) = +\infty$.

1.3 Moments and Related Problems

In this section we suppose μ to be such that

$$\text{Int}\,\Omega(\mu) = (\omega_0, \omega_1) \neq \phi ,$$

and (1.18)

$$\text{Int}\,\Lambda(\mu) \neq \phi .$$

The proof of the following theorem can be found in [4] and is indicated here.

Theorem 1.4. Suppose μ satisfies (1.18) and $F(\cdot\,;\omega)$ is the distribution corresponding to $\omega \in \text{Int}\,\Omega(\mu)$. Then all of the moments of $F(\cdot\,;\omega)$ exist, and $\frac{dm(\omega)}{d\omega}$ exists and is positive.

Proof: Let ω be such that $\omega_0 < \omega < \omega_1$ and let

$$T = \min\{\omega-\omega_0,\ \omega_1-\omega\} > 0 .$$

For all t such that $|t| < T$, $t + \omega \in \Omega(\mu)$. Therefore

$$0 < \int e^{(\omega+t)x} d\mu(x) < +\infty$$

which implies

$$\psi(t) = \beta(\omega) \int e^{(\omega+t)x} d\mu(x) < +\infty .$$

Thus $\psi(t)$ exists whenever $|t| < t$, and all moments of $F(\cdot\,;\omega)$ exist for $\omega_0 < \omega < \omega_1$.

Specifically

$$m(\omega) = \beta(\omega) \int x e^{\omega x} d\mu(x) < +\infty .$$

$$\frac{dm(\omega)}{d\omega} = n(\omega) \frac{\int x^2 e^{\omega x} d\mu(x) - \left(\int x e^{\omega x} d\mu(x)\right)^2}{[n(\omega)]^2}$$

$$= \int \frac{x^2 e^{\omega x}}{n(\omega)} d\mu(x) - [m(\omega)]^2$$

$$= \beta(\omega) \int (x - m(\omega))^2 e^{\omega x} d\mu(x)$$

$$= \sigma^2(\omega) > 0 .$$

It is intuitively clear that the mean of a distribution generated by μ must be contained in $\Lambda(\mu)$. The range of a function f is a set denoted by $R(f(\cdot))$.

Theorem 1.5. $R(m(\cdot)) \subset \Lambda(\mu)$.

Proof: By definition of b,

$$x \leq b \quad \text{a.e.} \quad [\mu].$$

Then

$$xe^{\omega x} \leq be^{\omega x} \quad \text{a.e.} \quad [\mu]$$

and

$$\beta(\omega) \int xe^{\omega x} d\mu(x) \leq \beta(\omega) \int be^{\omega x} d\mu(x) = b.$$

Thus $m(\omega) \leq b$. Similarly $m(\omega) \geq a$.

We have established the existence of a one-one continuous mapping of $\text{Int}\,\Omega(\mu)$ onto $R(m(\cdot)) \subset \Lambda(\mu)$. Therefore there exists a function $\omega(\cdot) = m^{-1}(\cdot)$ which is a one-one continuous mapping of $R(m(\cdot))$ onto $\text{Int}\,\Omega(\mu)$. We have

$$m(\omega(\lambda)) = \lambda, \quad \lambda \in R(m(\cdot)). \tag{1.19}$$

(1.19) is a very convenient equation since we can alter the parameter space to $R(m(\cdot))$ as shown in the following equation.

$$F(x;\omega(\lambda)) = F(x|\lambda) = \beta(\omega(\lambda)) \int_{(-\infty, x]} e^{\omega(\lambda)x} d\mu(x),$$

for $\lambda \in R(m(\cdot))$.

The parameter space corresponding to the distribution functions $F(\cdot|\lambda)$ is $R(m(\cdot))$. By Proposition 1.5, whenever

$\Lambda(\mu)$ is a bounded subset of R, $\Omega(\mu) = R$. Thus, the new parameter space must be a bounded set.

We now consider two important examples.

Example 1.4. Let m be Lebesgue measure restricted to R^+. $\Lambda(m) = [0, \infty)$.

$$n(\omega) = \int_0^{\infty} e^{\omega x} dx = \frac{-1}{\omega}$$

whenever $\omega \epsilon \Omega(m) = (-\infty, 0)$.

It is easily verified that

$$m(\omega) = \frac{1}{\omega^2} \left[\frac{-1}{\omega} \right]^{-1} = \frac{-1}{\omega} .$$

$m(\omega)$ has the desired properties and $\omega(\cdot)$ is defined by

$$\omega(\lambda) = \frac{-1}{\lambda} \quad \text{for} \quad \lambda \epsilon (0, \infty).$$

Thus

$$F(x|\lambda) = \int_0^x \frac{1}{\lambda} e^{-\frac{t}{\lambda}} dt, \quad x \geq 0 . \tag{1.20}$$

Example 1.5. Let ν be counting measure on $M = N \cup \{0\}$ (M will be used below). ($\nu(A)$ is defined to be the number of elements of M contained in A.) $\Lambda(\nu) = [0, \infty)$.

$$n(\omega) = \sum_{k=0}^{\infty} e^{\omega x} < +\infty$$

whenever $e^{\omega} < 1$ or $\omega \epsilon (-\infty, 0) = \Omega(\nu)$. Therefore

$$n(\omega) = \frac{1}{1-e^{\omega}} .$$

It can be verified that

$$m(\omega) = \frac{e^{\omega}}{1-e^{\omega}}$$

which has the properties listed above. $\omega(\lambda)$ is defined by

$$\omega(\lambda) = \ell n \frac{\lambda}{1-\lambda} .$$

Thus

$$F(x|\lambda) = \sum_{k=0}^{[x]} \left(\frac{\lambda}{1+\lambda}\right)^k \frac{1}{1+\lambda} , \quad x \geq 0 . \qquad (1.21)$$

In equation (1.20), $F(\cdot|\lambda)$ is the distribution function of an exponential random variable with parameter $1/\lambda$. In equation (1.21), $F(\cdot|\lambda)$ is a geometric distribution function with parameter $1/(1+\lambda)$. In each example $m(\cdot)$ accomplished a mapping of $\text{Int}\,\Omega$ <u>onto</u> $\text{Int}\,\Lambda$. There are examples (the binomial distribution, for instance) of the mean mapping the real line onto $\text{Int}\,\Lambda$, a bounded set.

We now consider the question of how often $m(\cdot)$ maps $\text{Int}\,\Omega$ onto $\text{Int}\,\Lambda$. In [4] a parameterization of exponential type distributions is given in which it is assumed that (1.19) holds for all $\lambda \in \Lambda(\mu)$. To explore the generality of this assumption we consider slightly weaker conditions. Our object is to characterize measures such that

$$R(m(\cdot)) \supset \text{Int}\,\Lambda(\mu). \tag{1.22}$$

We will examine the problem in detail, but we first consider a very restricted set of measures.

<u>Theorem 1.6.</u> Assume $\Lambda(\mu) = [a,b]$ is a bounded subset of R and that a and b are atoms of μ. Then $R(m(\cdot)) \supset \text{Int}\,\Lambda(\mu)$.

<u>Proof:</u> By the bounded convergence theorem

$$\lim_{k \to \infty} \int e^{k(x-b)} d\mu(x) = \mu(\{b\}),$$

and

$$\lim_{k \to \infty} \int \frac{x}{b} e^{k(x-b)} d\mu(x) = \mu(\{b\}).$$

Therefore

$$\lim_{k\to\infty} \frac{\int xe^{kx}d\mu(x)}{\int e^{kx}d\mu(x)} = \lim_{k\to\infty} \frac{be^{kb}\mu(\{b\})}{e^{kb}\mu(\{b\})} = b\ .$$

Similarly $\lim_{k\to-\infty} m(k) = a.$

Suppose, however, that $\mu(\{b\}) = 0$. Then we have

$$\lim_{k\to\infty} \int e^{k(x-b)}d\mu(x) = 0\ ,$$

and the method of the preceding theorem does not apply. To proceed we use an idea of Laplace (known as Laplace's method [2]) about the asymptotic expansion of integrals and apply it to the integrals under consideration. We will take

$$a(x) \sim b(x) \quad (x \to x_0)$$

to mean

$$\lim_{x\to x_0} \frac{a(x)}{b(x)} = 1\ .$$

Lemma. Suppose $\Lambda(\mu) = [a, b]$, $\Omega(\mu) \neq \phi$, $f \geq 0$, $f \in L_1(\mu)$ and $0 < \varepsilon < b-a$. Then, if $\int_{[b-\delta, b]} f d\mu > 0$ for all $\delta > 0$, we have

$$\int_{[a,b]} f(x)e^{kx}d\mu(x) \sim \int_{[b-\varepsilon,b]} f(x)e^{kx}d\mu(x). \qquad (k \to \infty)\ .$$

Proof:

$$\int_{[a,b]} f(x)e^{k(x-b+\varepsilon)}d\mu(x) = \int_{[a,b-\varepsilon]} f(x)e^{k(x-b+\varepsilon)}d\mu(x) + \int_{[b-\varepsilon,b]} f(x)e^{k(x-b+\varepsilon)}d\mu(x).$$

By the bounded convergence theorem,

$$\lim_{k\to\infty} \int_{[a,b-\varepsilon]} f(x)e^{k(x-b+\varepsilon)}d\mu(x) = 0.$$

Also, by the monotone convergence theorem,

$$\int_{[b-\varepsilon,b]} f(x)e^{k(x-b+\varepsilon)}d\mu(x)$$

becomes unbounded as $k \to \infty$. Thus we may conclude

$$\int_{[a,b]} f(x)e^{k(x-b+\varepsilon)}d\mu(x) \sim \int_{[b-\varepsilon,b]} f(x)e^{k(x-b+\varepsilon)}d\mu(x). \qquad (k \to \infty)$$

Whenever $a_n b_n \sim a_n c_n$, we have $b_n \sim c_n$.

Therefore, letting $a_k = e^{k(-b+\varepsilon)}$,

$$\int_{[a,b]} f(x)e^{kx}d\mu(x) \sim \int_{[b-\varepsilon,b]} f(x)e^{kx}d\mu(x).$$

The above lemma is extended to $\Lambda(\mu)$ unbounded in the lemma preceding Theorem 1.8.

As an application of such an expansion let $g(x) = \frac{d\mu}{dm}(x)$ be bounded on $[a, b]$, and left-continuous and non-zero at b. Then it can be shown that

$$\int_a^b g(x)e^{kx}dx \sim g(b)e^{kb}k^{-1}$$

and

$$\int_a^b xg(x)e^{kx}dx \sim bg(b)e^{kb}k^{-1} .$$

This, of course, implies $\lim_{k\to\infty} m(k) = b$.

<u>Theorem 1.7.</u> Let μ be a measure such that $\Lambda(\mu) = [a, b]$ and $\Omega(\mu) = R$. Then

$$\lim_{x\to\infty} m(x) = b. \tag{1.23}$$

and

$$\lim_{x\to-\infty} m(x) = a. \tag{1.24}$$

Proof: Note that the assumption $\Omega(\mu) = R$ is implied by $\Omega(\mu) \neq \phi$. If $a = b$, the conclusion is obvious. Thus we assume $\text{Int}\,\Lambda(\mu) \neq \phi$. We will prove (1. 23) by taking the limit of a sequence and then applying the continuity of $m(\cdot)$.

By an application of the lemma,

$$\lim_{k \to \infty} \frac{\int_{[a,b]} xe^{kx} d\mu(x)}{\int_{[a,b]} e^{kx} d\mu(x)} = \lim_{k \to \infty} \frac{\int_{[b-\varepsilon,b]} xe^{kx} d\mu(x)}{\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)} .$$

$$\lim_{k \to \infty} \frac{b\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)}{\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)} \geq \lim_{k \to \infty} \frac{\int_{[b-\varepsilon,b]} xe^{kx} d\mu(x)}{\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)} \geq \lim_{k \to \infty} \frac{(b-\varepsilon)\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)}{\int_{[b-\varepsilon,b]} e^{kx} d\mu(x)} .$$

Thus we may conclude $b \geq \lim_{k \to \infty} m(k) \geq b-\varepsilon$ for arbitrary $\varepsilon > 0$.

Therefore $\lim_{k \to \infty} m(k) = b$. (1. 24) follows in the same manner.

It should be noted that Theorem 1. 7 implies Theorem 1. 6 which is presented for its simple proof and motivation for new considerations. Thus (1. 22) is true for all measures with non-empty parameter spaces and bounded spectra.

Lemma. Suppose $\Lambda(\mu)$ is unbounded on the right and $\Omega(\mu) = R$. If $g \in L_1(\mu)$ is a non-negative function such that, for every $B < +\infty$,

$$\int_{[B,\infty)} g(x)\, d\mu(x) > 0,$$

then

$$\int g(x)e^{kx} d\mu(x) \sim \int_{[A,\infty)} g(x)e^{kx} d\mu(x) \qquad (k \to +\infty)$$

for all $A < +\infty$.

Proof:

$$\lim_{k \to \infty} \int_{(-\infty, A)} g(x)e^{k(x-A)} d\mu(x) = 0$$

and

$$\lim_{k \to \infty} \int_{[A,\infty)} g(x)e^{k(x-A)} d\mu(x) = +\infty.$$

Therefore

$$\int g(x)e^{k(x-A)} d\mu(x) \sim \int_{[A,\infty)} g(x)e^{k(x-A)} d\mu(x)$$

and

$$\int g(x)e^{kx} d\mu(x) \sim \int_{[A,\infty)} g(x)e^{kx} d\mu(x).$$

Theorem 1.8. Let μ be a measure such that $\Lambda(\mu)$ is unbounded on the right (that is, $b = +\infty$) and $\Omega(\mu) = R$. Then $R(m(\cdot)) \supset \text{Int}\,\Lambda(\mu)$.

Proof.

$$\lim_{k \to \infty} \frac{\int x e^{kx} d\mu(x)}{\int e^{kx} d\mu(x)} = \lim_{k \to \infty} \frac{\int_{[A,\infty)} x e^{kx} d\mu(x)}{\int_{[A,\infty)} e^{kx} d\mu(x)} \geq \lim_{k \to \infty} \frac{A \int_{[A,\infty)} e^{kx} d\mu(x)}{\int_{[A,\infty)} e^{kx} d\mu(x)} = A,$$

or $\lim_{k \to \infty} m(k) \geq A$ for all $A < +\infty$. Thus (1.23) holds.

If $a = -\infty$, then (1.24) follows by a similar argument. If $a > -\infty$, then we can apply the proof of Theorem 1.7 to yield (1.24).

The remaining cases are treated in the following theorem. We study only $\lim_{x \to \omega_1^-} m(x)$. However, the limits we do not treat follow from the theorems we present, and our treatment of the problem is complete with Theorem 1.9.

Theorem 1.9. Let μ be a measure such that $\Lambda(\mu)$ is unbounded on the right and $\text{Int}\,\Omega(\mu) = (\omega_0, \omega_1) \neq \phi$, where $\omega_1 < +\infty$. Then

(i) If $\omega_1 \notin \Omega(\mu)$,

$$\lim_{\omega \to \omega_1^-} m(\omega) = +\infty; \tag{1.25}$$

(ii) If $\omega_1 \in \Omega(\mu)$ but $xe^{\omega_1 x} \notin L_1(\mu)$,

$$\lim_{\omega \to \omega_1^-} m(\omega) = +\infty ; \tag{1.26}$$

(iii) If $\omega_1 \in \Omega(\mu)$ and $xe^{\omega_1 x} \in L_1(\mu)$,

$$\lim_{\omega \to \omega_1^-} m(\omega) < +\infty . \tag{1.27}$$

Note that (iii) implies that there are members of the exponential family such that $m(\cdot)$ maps $\text{Int}\,\Omega$ <u>into</u> $\text{Int}\,\Lambda$.

<u>Proof:</u> (i) If $\omega_1 \notin \Omega(\mu)$, we have, for $\omega \in \Omega(\mu)$,

$$\int_{R^-} e^{\omega_1 x} d\mu(x) \leq \int_{R^-} e^{\omega x} d\mu(x) < +\infty, \tag{1.28}$$

$$\int_{R^-} |x| e^{\omega_1 x} d\mu(x) \leq \int_{R^-} |x| e^{\omega x} d\mu(x) < +\infty , \tag{1.29}$$

and

$$\int_{x \geq 1} xe^{\omega_1 x} d\mu(x) \geq \int_{x \geq 1} e^{\omega_1 x} d\mu(x) = +\infty . \tag{1.30}$$

For $A \geq 1$ we have

$$\int e^{\omega x} d\mu(x) \sim \int_{[A,\infty)} e^{\omega x} d\mu(x) \qquad (\omega \to \omega_1^-)$$

by (1.28), and

$$\int xe^{\omega x} d\mu(x) \sim \int_{[A,\infty)} xe^{\omega x} d\mu(x) \qquad (\omega \to \omega_1^-)$$

by (1.29) and (1.30).

Thus

$$\lim_{\omega \to \omega_1^-} m(\omega) = \lim_{\omega \to \omega_1^-} \frac{\int_{[A,\infty)} xe^{\omega x} d\mu(x)}{\int_{[A,\infty)} e^{\omega x} d\mu(x)} \geq A,$$

and (1.25) follows.

(ii) By the monotone and bounded convergence theorems,

$$\lim_{\omega \to \omega_1^-} \frac{\int xe^{\omega x} d\mu(x)}{\int e^{\omega x} d\mu(x)} = \lim_{\omega \to \omega_1^-} \frac{\int xe^{\omega x} d\mu(x)}{\int e^{\omega_1 x} d\mu(x)} = +\infty .$$

(iii) Using the bounded convergence theorem,

$$\lim_{\omega \to \omega_1^-} \frac{\int xe^{\omega x} d\mu(x)}{\int e^{\omega x} d\mu(x)} = \frac{\int xe^{\omega_1 x} d\mu(x)}{\int e^{\omega_1 x} d\mu(x)} < +\infty.$$

Therefore (1.23) holds unless the situation described in (iii) of Theorem 1.9 occurs. Clearly, $\text{Int}\,\Lambda(\mu) = \phi$ implies $\Lambda(\mu) = [a, a]$ for some a, and $R(m(\cdot)) = \Lambda(\mu)$. Since (iii) is the only case in which (1.22) is false, we refer to Example 1.3 with $\delta > 1$ for an example of this situation. We summarize our results in the following theorem.

Theorem 1.10. Let μ be a measure such that $\text{Int}\,\Omega(\mu) = (\omega_0, \omega_1)$ is non-empty. Then

$$R(m(\cdot)) \supset \text{Int}\,\Lambda(\mu)$$

unless (i) $\omega_1 < +\infty$, and $xe^{\omega_1 x}$ is integrable;

or (ii) $-\infty < \omega_0$, and $xe^{\omega_0 x}$ is integrable.

1.4 Members of the Exponential Family

This section gives futher examples of distributions found in the exponential family. Propositions 1.2 and 1.3 dealt with improper distributions. Besides this, five examples have been given; two of these are catalogued below for completeness. One of the more frequently encountered distributions on the non-negative integers not found below is the hypergeometric. Indeed, it is not a member of the exponential family. Despite this omission many of the common probability distributions belong to the exponential family.

The first group of distributions will be generated by measures on M, the non-negative integers. If μ is such a measure, it will be assumed that

$$\mu(R-M) = 0.$$

The function $p(x; \omega)$ will represent the probability mass function of the distribution corresponding to ω.

(A) Suppose μ is a measure such that $\mu(\{a\}) > 0$ and $\mu(R-\{a\}) = 0$. Then μ determines an improper distribution with jump at $x = a$. (See equation (1.8).)

(B) Suppose μ is a measure such that

$$\mu(\{0\}) = 1,$$

$$\mu(\{1\}) = 1,$$

$$\mu(M-\{0,1\}) = 0.$$

$n(\omega) = 1 + e^{\omega}$ so that

$$p(0;\omega) = \frac{1}{1+e^{\omega}} ,$$

and

$$p(1;\omega) = \frac{e^{\omega}}{1+e^{\omega}} .$$

Clearly $p(\cdot\,;\omega)$ is the mass function of a Bernoulli random variable with probability of "success" equal to $\frac{e^{\omega}}{1+e^{\omega}}$.

(C) Define μ by

$$\mu(\{x\}) = \binom{n}{x}, \quad x \in \{0, 1, 2, \cdots, n\} ,$$

and

$$\mu(M - \{0, 1, 2, \cdots, n\}) = 0.$$

$$n(\omega) = \sum_{k=0}^{n} \binom{n}{k} e^{\omega k} = (1+e^{\omega})^{n} .$$

$$p(k;\omega) = \binom{n}{k} \left(\frac{e^{\omega}}{1+e^{\omega}}\right)^{k} \left(\frac{1}{1+e^{\omega}}\right)^{n-k} , \quad k = 0, 1, \cdots, n .$$

$p(\cdot\,;\omega)$ is the mass function of a binomial random variable with parameters n and $\frac{e^{\omega}}{1+e^{\omega}}$.

Of course (C) contains (B) but they are listed separately. This will also occur with (F) and (E) below and will not be mentioned again.

(D) Let μ be a measure such that

$$\mu(\{x\}) = \frac{1}{x!}, \quad x \in M .$$

$$n(\omega) = \sum_{k=0}^{\infty} \frac{e^{\omega k}}{k!} = e^{\lambda} \text{ where } \lambda = e^{\omega}. \quad \Omega(\mu) = R.$$

$$p(k;\omega) = \frac{e^{-\lambda}\lambda^{k}}{k!}, \quad k \in M$$

$p(\cdot\,;\omega)$ is the probability mass function of a Poisson random variable with mean e^{ω}.

(E) Assume μ is a measure with

$$\mu(\{x\}) = 1, \quad x \in M.$$

$$n(\omega) = (1-e^{\omega})^{-1}. \quad \Omega(\mu) = (-\infty, 0).$$

$$p(k;\omega) = (1-e^{\omega})e^{\omega k}.$$

$p(\cdot\,;\omega)$ is the mass function of a geometric probability law with parameter $1-e^{\omega}$. (See Example 1.6.)

(F) Define μ by

$$\mu(\{x\}) = \frac{(m+x-1)!}{x!}, \quad x \in M,$$

where $m \in N$.

$$n(\omega) = \sum_{k=0}^{\infty} e^{\omega k} \frac{(m+k-1)!}{k!} = (1-e^{\omega})^{-m}(m-1)!$$

whenever $\omega \in \Omega(\mu) = (-\infty, 0)$.

$$p(k;\omega) = \binom{m+k-1}{k} (1-e^{\omega})^{m} e^{\omega k}, \quad k \in M .$$

Therefore we have the negative binomial distribution with parameters m and $1-e^{\omega}$.

(G) Suppose

$$\mu(\{x\}) = 1, \quad x = 0, 1, \cdots, n,$$

and

$$\mu(M-\{0, 1, \cdots, n\}) = 0 .$$

This measure generates the truncated geometric distribution (see (D)). $\Omega(\mu) = R$ so that, in particular, $0 \in \Omega(\mu)$.

$$n(0) = \sum_{k=0}^{n} 1 = n+1$$

and

$$p(k;0) = \frac{1}{n+1}, \quad k = 0, 1, \cdots, n .$$

Thus we have the probability measure on $\{0, 1, 2, \cdots, n\}$ in which each point is equally likely.

(H) Assume

$$\mu(\{x\}) = \frac{1}{x}, \quad x \in N,$$

and

$$\mu(\{0\}) = 0 .$$

Then

$$n(\omega) = \sum_{k=1}^{\infty} \frac{e^{\omega k}}{k} = -\ell n(1-e^{\omega})$$

whenever $\omega \in \Omega(\mu) = (-\infty, 0)$.

$$p(k; \omega) = \frac{-1}{\ell n(1-e^{\omega})} \frac{e^{\omega k}}{k}, \quad k \in N.$$

$p(\cdot\,; \omega)$ is the mass function for Fisher's logarithmic series distribution.

Next we turn to the absolutely continuous distributions. m, as above, denotes Lebesgue measure, and we define measures on R by their Radon-Nikodym derivatives with respect to Lebesgue measure. The probability measures we generate will also be absolutely continuous and so will be characterized by their density functions $f(\cdot\,; \omega)$.

(I) Restrict m to $[0,\infty)$.

$$n(\omega) = -\frac{1}{\omega} \quad \text{for} \quad \omega \in \Omega(\mu) = (-\infty, 0).$$

$$f(x;\omega) = \begin{cases} -\omega e^{\omega x} = |\omega| e^{-|\omega| x}, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

$f(\cdot\,;\omega)$ is the density function of an exponential random variable with mean $|\omega|$. (See Example 1.5.)

(J) Define μ by

$$\frac{d\mu}{dm}(x) = \exp\{-\frac{x^2}{2\sigma^2}\}, \quad x \in R.$$

$$n(\omega) = \int \exp\{-\frac{x^2}{2\sigma^2} + \omega x\}\, dx$$

$$= \exp\{\frac{\omega^2\sigma^2}{2}\} \int \exp\{-\frac{1}{2}(\frac{x-\omega\sigma^2}{\sigma})^2\}\, dx$$

$$= \sigma\sqrt{2\pi}\, \exp\{\frac{\omega^2\sigma^2}{2}\}.$$

Therefore

$$f(x;\omega) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{\omega^2\sigma^2}{2} + \omega x - \frac{x^2}{2\sigma^2}\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2}(\frac{x-\omega\sigma^2}{\sigma})^2\},$$

and $f(\cdot\, ; \omega)$ is the density function of a normal random variable with mean $\omega\sigma^2$ and variance σ^2.

(K) Restrict m to $[a, b]$.

If $\omega \neq 0$,

$$n(\omega) = \int_a^b e^{\omega x} dx = \frac{1}{\omega} [e^{\omega b} - e^{\omega a}],$$

and

$$f(x; \omega) = \begin{cases} \dfrac{\omega}{e^{b\omega} - e^{a\omega}} e^{\omega x}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

When $a = 0$ we have the density function of a truncated exponential distribution. When $\omega = 0$, $n(0) = b-a$ and

$$f(x; 0) = \frac{1}{b-a}, \quad x \in [a, b].$$

Thus we have the uniform distribution on $[a, b]$.

(L) Define a measure μ by

$$\frac{d\mu}{dm}(x) = \begin{cases} x^{r-1}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

for $r > 0$.

$$n(\omega) = \int_0^\infty x^{r-1} e^{\omega x} dx = \frac{\Gamma(r)}{|\omega|^r} \int_0^\infty \frac{|\omega|^r x^{r-1} e^{-|\omega| x}}{\Gamma(r)} dx$$

$$= \Gamma(r) \, |\omega|^{-r},$$

whenever $\omega \epsilon \Omega(\mu) = (-\infty, 0)$.

$$f(x;\omega) = \frac{|\omega|^r}{\Gamma(r)} x^{r-1} e^{-|\omega| x}, \quad x > 0 ,$$

and is the density of a Gamma random variable with parameters r and $|\omega|$.

(M) Let μ be determined by

$$\frac{d\mu}{dm}(x) = \begin{cases} x^{\frac{n}{2}-1}, & x > 0 \\ 0, & x \leq 0 , \end{cases}$$

for $n \epsilon N$.

$$n(\omega) = \int_0^\infty x^{\frac{n}{2}-1} e^{\omega x} dx = 2^{\frac{n}{2}} \left(\frac{1}{\sqrt{2|\omega|}}\right)^n \Gamma\left(\frac{n}{2}\right)$$

for $\omega \epsilon \Omega(\mu) = (-\infty, 0)$.

$$f(x;\omega) = \frac{(|\omega|)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-|\omega| x}, \quad x > 0 .$$

Thus we have the density of a χ^2 random variable with parameters n and $\frac{1}{\sqrt{2|\omega|}}$.

From (K) and (G) it can be noted that any truncation of a distribution in the exponential family is again in the exponential family. It should not be supposed that these examples are shown here for the first time. Most of them are well known, yet do not appear in any collected form that this author could locate. Therefore it seemed worthwhile to compile a collection of examples and present them in the spirit of this discussion. Several sources have been helpful in this task [1, 3, 4, 7, 8].

CHAPTER 2. ATOMIC PROBABILITY DISTRIBUTIONS

The following definitions are taken from [3]. A point $x \in R$ is an atom of the measure μ if $\mu(\{x\}) > 0$. Let the distribution F be generated by μ. F is concentrated on the set A if $\mu(R-A) = 0$. F is said to be atomic if it is concentrated on the set of its atoms.

A well known atomic measure is counting measure on N. In this chapter we consider σ-finite atomic measures that are absolutely continuous with respect to counting measure on $M = N \cup \{0\}$. As seen in 1.4, many of the classical probability measures on M can be obtained from this class of measures. Throughout this chapter ν will denote counting measure on M.

2.1 Exponential Distributions on M

In this section we study exponential family distributions generated by a measure μ such that $\mu << \nu$ and $\mu(\{0\}) > 0$. A completely analogous treatment of absolutely continuous distributions exists and will be commented on below. Suppose we have such a measure. Since $\frac{d\mu}{d\nu}(x) = h(x)$ exists,

$$\int e^{\omega x} d\mu(x) = \sum_{k=0}^{\infty} e^{\omega k} h(k).$$

Therefore $\Omega(\mu) = \{\omega: 0 < \sum_{k=0}^{\infty} e^{\omega k} h(k) < +\infty\}$.

We define $n(\omega)$ and $\beta(\omega)$ as above and suppose that $\Omega(\mu) \neq \phi$. Of course, since $\Lambda(\mu) \subset [0, \infty)$, $\Omega(\mu)$ is unbounded on the left by Proposition 1.4.

For $x \in M$, $h(x) \geq 0$ and, moreover, $h(x) = \mu(\{x\})$. For our purposes, then, we need only consider non-negative functions $h(\cdot)$ on M such that $h(0) > 0$.

For examples consider (B) through (G) in 1.4. The function $h(x)$ has the form $\binom{n}{x}$, $\frac{1}{x!}$, 1, and $\frac{(m+x-1)!}{x!}$. The measures are exactly the same as in 1.4.

<u>Proposition 2.1.</u> Assume $h(\cdot)$ is defined on M and

(a) $h(k) \geq 0$, $k \in N$,

(b) $h(0) > 0$,

(c) $\Omega = \{\omega : \sum_{k=0}^{\infty} e^{\omega k} h(k) < +\infty\}$ is non-empty.

Then, for each $\omega \in \Omega$, $h(\cdot)$ determines a probability measure on M with positive weight on $\{0\}$. Furthermore, any function $g(\cdot)$ such that $g(k) = ah(k)$, $k \in M$, $a > 0$, determines the same measure.

Proof: Suppose $F(\cdot\,;\omega)$ is the distribution function associated with $h(\cdot)$ and $G(\cdot\,;\omega)$ the distribution function associated with $g(\cdot)$. Then

$$G(x;\omega) = \frac{\sum_{k=0}^{[x]} ah(k)e^{\omega k}}{\sum_{k=0}^{\infty} ah(k)e^{\omega k}} = F(x;\omega).$$

Proposition 2.1 implies a decomposition of the set of functions (or measures) satisfying (a), (b), and (c) into equivalence classes. Our equivalence relation " $\equiv$ " is, of course, defined by $f \equiv g$ if and only if there is an $a > 0$ such that $f = ag$. The functions within an equivalence class determine the same probability measure.

Assume that $F(\cdot\,;\omega)$ is the distribution function generated by $h(\cdot)$. Since $h(0) > 0$, we may, without loss of generality, assume $h(0) = 1$. (According to Proposition 2.1, $g(x) = \frac{h(x)}{h(0)}$ generates the same distribution as $h(x)$.) For $x \in M$

$$f(x) = \beta(\omega)\, h(x)e^{\omega x}, \quad \omega \in \Omega\,.$$

Substituting $x = 0$, we obtain

$$f(0) = \beta(\omega)h(0) = \beta(\omega)$$

or

$$f(0) = [\sum_{x=0}^{\infty} h(x)e^{\omega x}]^{-1} .$$

Now

$$f(x) = f(0)h(x)e^{\omega x} .$$

Clearly

$$\frac{f(0)f(x+y)}{f(x)\ f(y)} = \frac{h(x+y)}{h(x)\ h(y)} .$$

We have shown sufficiency in Theorem 2.1 below. The proof of necessity is due to Patil and Seshadri [7] .

Theorem 2.1. Let $f(\cdot)$ be the density function of an atomic probability distribution on M such that $f(0) > 0$. Then the distribution function of $f(\cdot)$ belongs to the exponential family if and only if there exists a non-negative function $h(\cdot)$ on M such that

$$\frac{f(0)\ f(x+y)}{f(x)\ f(y)} = \frac{h(x+y)}{h(x)\ h(y)} \tag{2.1}$$

whenever $f(x) > 0$ and $f(y) > 0$.

Proof: To complete the proof let

$$U(x) = \frac{f(x)}{f(0)h(x)} \tag{2.2}$$

Then

$$U(x+y) = U(x)\,U(y) \tag{2.3}$$

which is Cauchy's equation. Since $U(0) = 1$, (2.3) has a nontrivial solution and $U(x) = e^{\omega x}$ for some constant ω. Thus

$$f(x) = f(0)h(x)e^{\omega x}.$$

Letting μ be a measure such that $\mu \ll \nu$ and $\mu(\{x\}) = h(x)$, we have shown that f is a member of the exponential family.

The point of view in [7] is different from the one presented here and is enhanced by equation (2.4) below.

Suppose X and Y are independent random variables with atomic distributions and denote the conditional distribution of X given $X+Y$ by $c(x, x+y)$. Then

$$\frac{c(x+y,\ x+y)\,c(0,y)}{c(x,\ x+y)\,c(y,\ y)} = \frac{f(0)\,f(x+y)}{f(x)\ \ f(y)} \tag{2.4}$$

where f is the density of X.

From a set of assumptions involving X and Y, it is concluded that

$$f(x) = f(0)h(x)\, e^{\omega x}.$$

Theorem 2.1 has an immediate generalization. Let $\operatorname{supp}(\mu) = \{x: \mu(U) > 0$ for all open U containing $x\}$. Measures absolutely continuous with respect to μ are characterized by non-negative functions $h(\cdot)$ on $\operatorname{supp}(\mu)$. Suppose $f(\cdot)$ is the density of some probability measure with respect to μ and $\lambda \in \operatorname{supp}(\mu)$ is such that $f(\lambda) > 0$. Then f is the density of an exponential family distribution if and only if there exists a non-negative function $h(\cdot)$ defined on $\operatorname{supp}(\mu)$ such that

$$\frac{f(x+y)f(\lambda)}{f(x)\,f(y)} = \frac{h(x+y)}{h(x)h(y)}$$

whenever $f(x) > 0$ and $f(y) > 0$.

2.2 A Class of Distributions on M

The class of distribution functions, F_1, defined below appears in [4], and we define F_1 here for completeness. F_1 provides a very good illustration of the concepts in Chapter 1 and some observations will be made.

We define a class of measures on M by their Radon-Nikodym derivatives with respect to counting measure. Let μ be a measure on M such that

$$\frac{d\mu}{d\nu}(x) = \begin{cases} 1, & x = 0, \\ \dfrac{(\alpha+\beta)\cdots(\alpha+(x-1)\beta)}{x!}, & x \in N, \end{cases} \qquad (2.5)$$

where α and β are numbers such that (i) $\alpha \geq 0$ and $\beta \geq 0$ or (ii) $\alpha > 0$ and $\beta = -\frac{\alpha}{b}$ for some $b \in N$. Clearly, for case (i), $\Lambda(\mu) = [0,\infty)$ and, for case (ii), $\Lambda(\mu) = [0, b]$. Now define F_1 to be the set of exponential family distribution functions generated by the measures defined above. The following theorem is taken from [4].

Theorem 2.2. Assume that μ satisfies (2.5) with $\alpha > 0$. Then $\omega(\lambda)$ is given by

$$\omega(\lambda) = \ell n \frac{\lambda}{\alpha+\beta\lambda}. \qquad (2.6)$$

Furthermore $R(m(\cdot)) \supset \operatorname{Int} \Lambda(\mu)$.

The proof follows from a consideration of the generating function

$$\sum_{x=0}^{b} r(x)t^{x} = \begin{cases} e^{\alpha t} & \beta = 0, \\ (1-\beta t)^{-\frac{\alpha}{\beta}}, & \beta \neq 0. \end{cases} \qquad (2.7)$$

When $b = +\infty$, $\Omega(\mu) = (-\infty, \ell n \beta^{-1})$. Thus Theorem 1.10 implies $R(m(\cdot)) \supset \mathrm{Int}\,\Lambda(\mu)$.

It can be noted that $f(k) = \beta(\omega) \left(\frac{\lambda}{\alpha+\beta\lambda}\right)^k \frac{d\mu}{d\nu}(k)$ and

$$f(0) = \left[\sum_{k=0}^{\infty} \left(\frac{\lambda}{\alpha+\beta\lambda}\right)^k \frac{d\mu}{d\nu}(k)\right]^{-1} .$$

F_1 is defined as a subclass of the exponential family. The next proposition relates F_1 to Theorem 2.1.

<u>Proposition 2.2.</u> F_1 is a subclass of the distributions described in Theorem 2.1 with

$$h(x) = \frac{d\mu}{d\nu}(x) .$$

<u>Proof:</u>

$$\frac{f(0)f(x+y)}{f(x)\,f(y)} = \frac{\frac{d\mu}{d\nu}(x+y)}{\frac{d\mu}{d\nu}(x)\,\frac{d\mu}{d\nu}(y)} ,$$

and (2.1) is satisfied.

<u>Examples.</u>

(a) Let $\alpha = 1$ and $\beta = -1$. Then

$$e^{\omega(\lambda)} = \frac{\lambda}{1-\lambda} .$$

The corresponding distribution is of a Bernoulli random variable having the value 1 with probability λ.

(b) Let $\alpha = m$ and $\beta = -1$. Then

$$e^{\omega(\lambda)} = \frac{\lambda}{m-\lambda} .$$

We obtain the distribution function of a binomial random variable with parameters m and $\frac{\lambda}{m}$.

(c) Let $\alpha = 1$ and $\beta = 0$. Then

$$e^{\omega(\lambda)} = \lambda .$$

We get the mass function for a Poisson random variable with mean λ.

(d) Let $\alpha = r$ and $\beta = 1$ when $r \in N$.

$$e^{\omega(\lambda)} = \frac{\lambda}{r+\lambda}.$$

$$p(x|\lambda) = \binom{r+x-1}{x} \left(\frac{r}{r+\lambda}\right)^r \left(\frac{\lambda}{r+\lambda}\right)^x, \quad x \in N;$$

which is the mass function of a negative binomial random variable.

From the many examples of distributions belonging to F_1 it might be supposed that all the distributions described in Theorem 2.1 belong to F_1. The following example illustrates that this is false.

Example 2.1. Define a distribution $F(x; \omega)$ on M by

$$h(x) = \begin{cases} 1, & x = 0, 1, 2, \cdots, n, \\ 0, & x \geq n+1 \quad . \end{cases}$$

This clearly is a distribution in the exponential family. (See Example (G).) However, if $n \geq 2$ we have enough conditions to determine α and β from the equations

$$h(x) = \frac{d\mu}{d\nu}(x), \quad x = 0, 1, \cdots, n \, .$$

This yields $\alpha = \beta = 1$. Therefore, if $F(x; \omega)$ belongs to F_1, it must be a geometric distribution. However, it is the truncated geometric. Thus F_1 is a proper subclass of the distributions in Theorem 2.1.

It is also possible, by making $h(\cdot)$ sufficiently irregular, to construct distributions in the exponential family that are not even truncations of distributions in F_1.

BIBLIOGRAPHY

1. Blackwell, David and M. A. Girshick. Theory of games and statistical decisions. New York, Wiley, 1954. 355 p.

2. Erdelyi, A. Asymptotic expansions. Dover Publications, 1956. 108 p.

3. Feller, William. An introduction to probability theory and its applications. Vol. II. New York, Wiley, 1966. 626 p.

4. Guthrie, D., Jr. and M. V. Johns, Jr. Bayes acceptance sampling procedures for large lots. The Annals of Mathematical Statistics 30: 896-925. 1959.

5. Halmos, Paul R. Measure theory. Princeton, Van Nostrand, 1964. 304 p.

6. Lehmann, E. L. Testing statistical hypotheses. New York, Wiley, 1959. 369 p.

7. Patil, G. P. and V. Seshadri. Characterization theorems for some univariate probability distributions. Journal of the Royal Statistical Society, Series B, 26: 286-292. 1964.

8. Patil, G. P. and Richard Shorrock. On certain properties of the exponential-type families. Journal of the Royal Statistical Society, Series B, 27: 94-99. 1965.