Charles University in Prague
Faculty of Mathematics and Physics

BACHELOR THESIS

Martin Schenk

Characterization of probability distributions

Department of probability and mathematical statistics

Supervisor: Prof. Lev Klebanov
Study program: Mathematics
Study plan: Theory of probability and random processes

2006

Here I would like to thank my supervisor Prof. Lev Klebanov for his invaluable help during the whole process of writing this thesis. I also want to thank my parents and all who help me with my studies and thus contributed to the creation of this thesis.

I proclaim that I wrote my bachelor thesis singly and only with the use of the cited sources of information. I agree with lending of this thesis and it's exposure.

In Beroun on the 25^{th} of May

Martin Schenk

Název práce: Charakterizace pravděpodobnostních rozdělení
Autor: Martin Schenk
Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: Prof. Lev Klebanov
e-mail vedoucího: levkleb@yahoo.com

Abstrakt: Tato práce je věnována dvěma známým charakterizačním větám: Bernsteinově a Polyově, vztahu mezi nimi a jejich modifikaci. Nejprve dokáži tyto dvě věty, poté ukáži spojitost mezi nimi, která vyplyne ze způsobu důkazu obou vět. Hlavním nástrojem v obou důkazech budou charakteristické funkce a jejich základní vlastnosti. V další části použiji pokročilejší teorii charakteristických funkcí k důkazu zobecněné verze Bernsteinovy a Polyovy věty. Na konci své práce zmiňuji několik aplikací Polyovy věty.

Klíčová slova: analytické characteristické funkce, Cramérova věta, faktorizace charakteristických funkcí

Title: Characterization of probability distributions
Author: Martin Schenk
Department: Department of probability and mathematical statistics
Supervisor: Prof. Lev Klebanov
Supervisor's e-mail address: levkleb@yahoo.com

Abstract: This thesis is dedicated to two already known characterization theorems: Bernstein's theorem and Polya's theorem, to the relation between them and to the modification of this two theorems. Foremost I prove these two theorems, then I show the connection between them, which is evident from the way of their proof. The main tool I use in the two proofs are characteristic functions and their basic qualities. In the next part, I use more advanced theory of characteristic functions to prove more general version of both Bernstein's and Polya's theorem. At the end I give some applications of Polya's theorem.

Keywords: analytic characteristic functions, Cramer's theorem, factorization of characteristic functions

Contents

Chapter 1

Introduction

The characterization of probability distributions is a section of probability theory and mathematical statistics. It's corner stone is the monograph KAGAN A.M., LINNIK YU.V., RAO C.R. - *Characterization problems of mathematical statistics*, that is especially devoted to characterization theorems, which deal with subsequent problem: we have some set of identically distributed random variables with some desired property and want to find out the distribution of these random variables. From one point of view, we can say that there are two possibilities how to understand characterization theorems: either we want to find out some methods of proofs of characterization theorems, or we concentrate on certain concrete characterization theorems. I chose the second approach, concretely I will work with the Bernstein's theorem and the Polya's theorem, which give us some information about normally distributed random variables.

Chapter 2

Bernstein's theorem

Let X_1, X_2 be two independent random variables. We denote by

$$\begin{aligned} L_1 &= X_1 + X_2 \\ L_2 &= X_1 - X_2. \end{aligned}$$

Then L_1 and L_2 are independent $\iff$ X_1 and X_2 are normally distributed, $X_1 \sim \mathrm{N}(\mu_1, \sigma^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma^2)$.

Proof:

1. Let L_1 and L_2 be independent. Further on, we will denote by $f_Y(t)$ the value of the characteristic function of the random variable Y in the point t. From the independency of L_1 and L_2 we see that

   $$f_{(L_1,L_2)}(t,s) = f_{L_1}(t) \cdot f_{L_2}(s)$$

   Moreover

   $$\begin{aligned} f_{(L_1,L_2)}(t,s) &= \mathbb{E}e^{itL_1+isL_2} = \mathbb{E}e^{it(X_1+X_2)+is(X_1-X_2)} = \\ &= \mathbb{E}e^{iX_1(t+s)+iX_2(t-s)} = \mathbb{E}e^{iX_1(t+s)}e^{iX_2(t-s)} = \\ &= f_{X_1}(t+s) \cdot f_{X_2}(t-s) \end{aligned}$$

   from the independency of X_1 and X_2. Taking together the last two identities, we have

   $$f_{X_1}(t+s) \cdot f_{X_2}(t-s) = f_{L_1}(t) \cdot f_{L_2}(s)$$

   If we put $s = 0$, then $f_{L_1}(t) = f_{X_1}(t) \cdot f_{X_2}(t)$, analogously $f_{L_2}(s) = f_{X_1}(s) \cdot f_{X_2}(-s)$, together we get

   $$f_{X_1}(t+s) \cdot f_{X_2}(t-s) = f_{X_1}(t) \cdot f_{X_2}(t) \cdot f_{X_1}(s) \cdot f_{X_2}(-s).$$

We will now work with the functions $g_j = \log f_j$ there, where $f_j \neq 0$. Then

$$g_1(t+s) + g_2(t-s) = g_1(t) + g_2(t) + g_1(s) + g_2(-s).$$

Suppose that the derivatives g_j', g_j'' exist. If we differentiate this equation with respect to t and then with respect to s (in the first case we take s as a parameter, in the second t as a parameter), we get

$$\begin{aligned} g_1'(t+s) + g_2'(t-s) &= g_1'(t) + g_2'(t) \\ g_1''(t+s) + g_2''(t-s) &= 0. \end{aligned}$$

If we denote $u := t - s$, then the last equation can be written in the form

$$g_1''(u+2s) = g_2''(u) \qquad \text{for any } s \in \mathbb{R}, \text{where } f_{x_j} \neq 0.$$

Therefore

$$g_1'' \equiv g_2'' \equiv a \in \mathbb{C}$$

on the whole real line (except those points where $f_{x_j}(s) = 0$). Then

$$g_j(t) = at^2 + b_j t + c_j, \qquad b_j, c_j \in \mathbb{C}$$

$$f_{x_j}(t) = \exp\{at^2 + b_j t + c_j\}$$

whence we see that the functions g_j are well defined on the whole real line. Moreover,

$$|f_{x_j}(t)| \leq 1 \Longrightarrow a < 0, a := -\frac{\sigma^2}{2}; \qquad f(0) = 1 \Longrightarrow c_j = 0;$$

also $f_{x_j}(t) = \overline{f_{x_j}(-t)}$, therefore $b_j = i\mu_j$, $\mu_j \in \mathbb{R}$ and

$$f_{x_j}(t) = \exp\{i\mu_j t - \frac{\sigma^2 t^2}{2}\},$$

which are the characteristic functions of normally distributed random variables with $X_1 \sim \mathrm{N}(\mu_1, \sigma^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma^2)$.

2. Now let X_1, X_2 be two independent random variables, $X_1 \sim \mathrm{N}(\mu_1, \sigma^2)$, $X_2 \sim \mathrm{N}(\mu_2, \sigma^2)$. Then $X_1 + X_2 \sim \mathrm{N}(\mu_1 + \mu_2, 2\sigma^2)$ and $X_1 - X_2 \sim \mathrm{N}(\mu_1 - \mu_2, 2\sigma^2)$ according to convolution theorem. So we will compare

$$\begin{aligned} f_{(L_1,L_2)}(t,s) &= \mathbb{E}e^{itL_1 + isL_2} \qquad \text{and} \\ f_{L_1}(t) \cdot f_{L_2}(s) &= \mathbb{E}e^{itL_1} \cdot \mathbb{E}e^{itL_1} \end{aligned}$$

(a)

$$\begin{aligned}\mathbb{E}e^{itL_1+isL_2} &= \mathbb{E}e^{it(X_1+X_2)+is(X_1-X_2)} = \mathbb{E}e^{iX_1(t+s)+iX_2(t-s)} = \\ &= \mathbb{E}e^{iX_1(t+s)}e^{iX_2(t-s)} = \mathbb{E}e^{iX_1(t+s)} \cdot \mathbb{E}e^{iX_2(t-s)},\end{aligned}$$

because X_1 and X_2 are independent and therefore also $e^{iX_1(t+s)}$ and $e^{iX_2(t-s)}$ are independent. Further,

$$\mathbb{E}e^{iX_1(t+s)} \cdot \mathbb{E}e^{iX_2(t-s)} = f_{X_1}(t+s) \cdot f_{X_2}(t-s) =$$

$$= e^{i\mu_1(t+s)-\frac{(t+s)^2\sigma^2}{2}} e^{i\mu_2(t-s)-\frac{(t-s)^2\sigma^2}{2}} = e^{i\mu_1(t+s)+i\mu_2(t-s)}.$$

$$\cdot e^{-\frac{\sigma^2}{2}\left(t^2+2ts+s^2+t^2-2ts+s^2\right)} = e^{i\mu_1(t+s)+i\mu_2(t-s)-\sigma^2\left(t^2+s^2\right)}$$

(b)

$$\mathbb{E}e^{itL_1} \cdot \mathbb{E}e^{itL_1} = f_{L_1}(t) \cdot f_{L_2}(s) = e^{i(\mu_1+\mu_2)t-\frac{2\sigma^2t^2}{2}}.$$

$$\cdot e^{i(\mu_1-\mu_2)s-\frac{2\sigma^2s^2}{2}} = e^{i\mu_1(t+s)+i\mu_2(t-s)-\sigma^2\left(t^2+s^2\right)}$$

We see that (a) and (b) are identical, therefore L_1 and L_2 are independent.

Chapter 3

Polya's theorem

Let X_1, X_2 be two independent identically distributed random variables, whose characteristic function is symmetric. We denote by

$$\begin{aligned} L_1 &= X_1 \\ L_2 &= \frac{X_1 + X_2}{\sqrt{2}}. \end{aligned}$$

Then L_1 and L_2 are identically distributed $\iff$ X_1 (and therefore also X_2) is normally distributed, $X_1 \sim \mathrm{N}(0, \sigma^2)$.

Proof:

1. Let $X_1 \sim \mathrm{N}(0, \sigma^2)$, then from the independency and identical distribution of X_1 and X_2 we have

   $$\begin{aligned} f_{L_2}(t) &= \mathbb{E}e^{it\left(\frac{X_1+X_2}{\sqrt{2}}\right)} = \mathbb{E}e^{iX_1\left(\frac{t}{\sqrt{2}}\right)} \cdot e^{iX_2\left(\frac{t}{\sqrt{2}}\right)} = \mathbb{E}e^{iX_1\left(\frac{t}{\sqrt{2}}\right)} \cdot \mathbb{E}e^{iX_2\left(\frac{t}{\sqrt{2}}\right)} = \\ &= f^2_{X_1}\left(\frac{t}{\sqrt{2}}\right). \end{aligned}$$

   Since $X_1 \sim \mathrm{N}(0, \sigma^2)$,

   $$f_{L_2}(t) = f^2_{X_1}\left(\frac{t}{\sqrt{2}}\right) = e^{2(-\frac{\sigma^2}{2}(\frac{t}{\sqrt{2}})^2)} = e^{-\frac{\sigma^2 t^2}{2}} = f_{X_1}(t) = f_{L_1}(t),$$

   so L_1 and L_2 have the same characteristic function. Because the distribution functions and corresponding characteristic functions are in reciprocal unambiguous relation, L_1 and L_2 are identically distributed.

2. Let L_1 and L_2 have the same distribution function. From the previous part of the proof we have

   $$f_{L_2}(t) = f^2_{X_1}\left(\frac{t}{\sqrt{2}}\right)$$

and this must be equal to $f_{X_1}(t)$. We must now prove that $f_{X_1}(t)$ is a characteristic function of $\mathrm{N}(0,\sigma^2)$.

(a) $f_{X_1}(t)$ does not attend 0: if there would exist $t \in \mathbb{R}$, for which $f_{X_1}(t) = 0$, then also $f^2_{X_1}(\frac{t}{\sqrt{2}}) = 0$ and $f_{X_1}(\frac{t}{\sqrt{2}}) = 0$. But then

$$\forall n \in \mathbb{N} \qquad f_{X_1}(\frac{t}{2^{\frac{n}{2}}}) = 0,$$

from which follows that

$$\lim_{n\to\infty} f_{X_1}(\frac{t}{2^{\frac{n}{2}}}) = 0,$$

which is equal to $f_{X_1}(0)$ from the continuity of f_{X_1}, because every characteristic function is continuous. But the equality $f_{X_1}(0) = = 0$ is in contradiction with the fact that f_{X_1} is a characteristic function, for which must be valid $f_{X_1}(0) = 1$.

(b) due to (a) we can write $\Phi(t) := \log f_{X_1}(t)$ on the whole real line. Therefore

$$\Phi(t) = 2\Phi(\frac{t}{\sqrt{2}}).$$

For $t > 0$ we will now introduce the function $\Psi(t)$,

$$\Phi(t) = t^2\Psi(t).$$

The condition for $f_{X_1}(t)$ can therefore be written in the form

$$t^2\Psi(t) = 2\frac{t^2}{2}\Psi(\frac{t}{\sqrt{2}}) \qquad \text{or} \qquad \Psi(t) = \Psi(\frac{t}{\sqrt{2}}).$$

Since $t > 0$, there exists $\alpha \in \mathbb{R}$ such that $t = e^\alpha$. So we will re-write the last equation subsequently:

$$\Psi(e^\alpha) = \Psi(e^{\alpha - \frac{1}{2}\log 2}).$$

If we denote by $U(\alpha) = \Psi(e^\alpha)$, we get

$$U(\alpha) = U(\alpha - \frac{1}{2}\log 2),$$

so that $U(t)$ is a $\frac{1}{2}\log 2$-periodic function. Since $\Psi(t)=U(\log t)$ and $t^2\Psi(t) = \Phi(t)$, we have

$$\Phi(t) = t^2U(\log t) \qquad \text{and} \qquad f_{X_1}(t) = \exp\{t^2U(\log t)\}.$$

Because $\lim_{x\to 0}\frac{1-e^x}{x}=1$, we can write

$$1 - f_{X_1}(t) \sim t^2 U(\log t) \qquad \text{or} \qquad \frac{1-f_{X_1}(t)}{t^2} \sim U(\log t)$$

for $t \to 0_+$. If there would be

$$\lim_{t\to 0_+} \frac{1-f(t)}{t^2} = +\infty,$$

(the limit must exist because of the periodicity of $U(t)$), then there must exist some $t_1 \in \mathbb{R}$, $t_1 > 0$ such that $U(\log t) = +\infty$, but then also

$$f_{X_1}(t_1) = \exp\{t^2 U(\log t)\} = +\infty,$$

which is in contradiction with the fact, that for every characteristic function is valid: $|f_{X_1}(t)| \leq 1 \ \forall t \in \mathbb{R}$. Consequently,

$$\lim_{t\to 0_+} \frac{1-f(t)}{t^2} < +\infty.$$

Because $f_{X_1}(t)$ is symmetric, we can write

$$\begin{aligned} 1 - f_{X_1}(t) &= 1 - \int_{-\infty}^{\infty} \cos tx \; dF(x) = \int_{-\infty}^{\infty} 1 - \cos tx \; dF(x) = \\ &= 2\int_{-\infty}^{\infty} \sin^2 \frac{tx}{2} \; dF(x) \end{aligned}$$

and therefore

$$\lim_{t\to 0_+} \frac{1-f(t)}{t^2} = \lim_{t\to 0_+} \frac{1}{2} \int_{-\infty}^{\infty} \frac{\sin^2 \frac{tx}{2}}{(\frac{tx}{2})^2} x^2 \; dF(x),$$

if both limits exist. We know from previous that $\lim_{x\to 0_+} \frac{1-f(t)}{t^2}$ is bounded. Moreover,

$$\liminf_{t\to 0_+} \frac{1-f(t)}{t^2} \geq \frac{1}{2} \liminf_{t\to 0_+} \int_{-A}^{A} \frac{\sin^2 \frac{tx}{2}}{(\frac{tx}{2})^2} x^2 \; dF(x)$$

for any $A > 0, A \in \mathbb{R}$. The function

$$\frac{\sin^2 \frac{tx}{2}}{(\frac{tx}{2})^2} x^2$$

is non-negative and

$$\int_{-A}^{A} \frac{\sin^2 \frac{tx}{2}}{(\frac{tx}{2})^2} x^2 \, dF(x) < \infty,$$

so we move the limit under the integral sign and get

$$\infty > \liminf_{t \to 0_+} \frac{1 - f(t)}{t^2} \geq \frac{1}{2} \int_{-A}^{A} x^2 \, dF(x).$$

If we apply $\lim_{A \to \infty}$ on both sides of this inequality, we see that $\int_{-\infty}^{\infty} x^2 \, dF(x)$ is finite and X_1 has finite second moment. We know that if any random variable has finite k-th moment, then there exist it's characteristic function's derivatives up to order k and these are continuous (this theorem can be found in [1]). So in our case f_{X_1} has continuous derivatives up to order 2, therefore also $\Phi''(t)$ is continuous, especially in 0, because

$$\Phi''(t) = \frac{f''(t)f(t) - (f'(t))^2}{(f(t))^2},$$

where $f(t), f'(t)$ and $f''(t)$ are continuous in 0. Then the equation

$$\Phi(t) = 2\Phi(\frac{t}{\sqrt{2}})$$

can be differentiated twice and we get

$$\Phi''(t) = \Phi''(\frac{t}{\sqrt{2}}).$$

Thus $\Phi''(t)$ is a constant: let us presume that there exist two real numbers $t, s > 0$ such that

$$\Phi''(t) = c_1, \Phi''(s) = c_2, \qquad c_1, c_2 \in \mathbb{C}, s \neq 2kt, k \in \mathbb{Z}.$$

Then from the continuity of $\Phi''(t)$ in zero follows that

$$\begin{aligned} \Phi''(0) &= \lim_{n \to \infty} \Phi''(\frac{t}{2^{\frac{n}{2}}}) = c_1 \qquad \text{but also} \\ \Phi''(0) &= \lim_{n \to \infty} \Phi''(\frac{s}{2^{\frac{n}{2}}}) = c_2 \end{aligned}$$

which leads to contrary.

(c) due to (b) we have $\Phi''(t) = a \ \forall t \in \mathbb{R},\ a \in \mathbb{C}$ and

$$\begin{aligned} \Phi(t) &= at^2 + bt + c, \qquad b, c \in \mathbb{C} \\ f(t) &= \exp\{at^2 + bt + c\}. \end{aligned}$$

Now we apply the neccesary conditions for a characteristic function:

$$\begin{aligned} f(0) = 1 &\implies c = 0 \\ f(t) = \overline{f(-t)} &\implies a \in \mathbb{R},\ b = i\mu,\ \mu \in \mathbb{R} \\ |f(t)| &\implies a < 0, \end{aligned}$$

we denote $a = -\frac{1}{2}\sigma^2$. Finally, $f(t) = f^2(\frac{t}{\sqrt{2}})$, therefore $\mu = 0$ and

$$f(t) = \exp\{-\frac{1}{2}\sigma^2 t^2\},$$

which is a characteristic function of $\mathrm{N}(0, \sigma^2)$.

Chapter 4

Connection between Bernstein's and Polya's theorem

These two theorems look quite different on the first sight. One concerns the independency of two linear combinations of independent random variables, the second studies the distribution of other linear combination of two independent random variables, which are in this case identically distributed. The independency and distribution of random variables are two quite different concepts, but as is clear from the proofs of this two theorems, they are closely linked through characteristic functions, especially by two key qualities of characteristic functions:

1. the distribution function and the corresponding characteristic function are in reciprocal unambiguous relation
2. for independent random variables is valid: the sum of two independent random variables has the characteristic function equal to the product of the characteristic functions of the two random variables.

We can therefore re-write our two theorems subsequently:
Bernstein's theorem:

$$f_{X_1}(t+s)\cdot f_{X_2}(t-s) = f_{L_1}(t)\cdot f_{L_2}(s) \iff \begin{aligned} f_{x_1}(t) &= \exp\{i\mu_1 t - \frac{\sigma^2 t^2}{2}\} \\ f_{x_2}(t) &= \exp\{i\mu_2 t - \frac{\sigma^2 t^2}{2}\} \end{aligned}$$

Polya's theorem:

$$f_{X_1}(t) = f_{X_1}^2(\frac{t}{\sqrt{2}}) \iff f_{x_1}(t) = \exp\{-\frac{\sigma^2 t^2}{2}\},$$

where in both theorems we have $a_1, a_2 \in \mathbb{C}, \sigma^2 > 0$. So we see that these are some kind of functional equations for characteristic functions. If the two random variables would be identically distributed also in the Bernstein's theorem, the close connection between these two theorems would be even clearer, because the functional equations would have consequential form:

$$\begin{aligned} f_{X_1}(t+s)f_{X_1}(t-s) &= f_{X_1}(t)f_{X_1}(t)f_{X_1}(s)f_{X_1}(-s), \\ f_{X_1}(2t) &= f^3_{X_1}(t)f_{X_1}(-t) \qquad \text{for } t = s \end{aligned}$$

while in Polya's theorem

$$f_{X_1}(2t) = f^4_{X_1}(t).$$

Chapter 5

Modification of Bernstein's and Polya's theorem

Let us see now, how the situation would change, if we take not two but four random variables in each theorem. The modification of Polya's theorem can be proved easily from the original theorem, because the functional equation

$$f(t) = f^2(\frac{t}{\sqrt{2}})$$

is equivalent with

$$f(t) = f^{2^n}(\frac{t}{2^{\frac{n}{2}}}),$$

therefore also for any $n \in \mathbb{N}$

$$X_1 \stackrel{d}{=} \frac{\sum_{i=1}^{2^n} X_i}{2^{\frac{n}{2}}} \iff X_1 \sim \mathrm{N}(0, \sigma^2)$$

(of course if X_i are iid).

In the modification of the Bernstein's theorem we will have

$$\begin{aligned} L_3 &= (X_1 + X_2) + (X_3 + X_4) \\ L_4 &= (X_1 + X_2) - (X_3 + X_4). \end{aligned}$$

We denote by

$$\begin{aligned} Y_1 &:= (X_1 + X_2) \\ Y_2 &:= (X_3 + X_4). \end{aligned}$$

Clearly, if $X_1 \sim \mathrm{N}(a_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(a_2, \sigma_2^2)$, $X_3 \sim \mathrm{N}(a_3, \sigma_3^2)$, $X_4 \sim \mathrm{N}(a_4, \sigma_4^2)$, where $\sigma_1^2 + \sigma_2^2 = \sigma_0^2 = \sigma_3^2 + \sigma_4^2$, then (according to convolution theorem) we will have $Y_1 \sim \mathrm{N}(a_1 + a_2, 2\sigma_0^2)$,$Y_2 \sim \mathrm{N}(a_3 + a_4, 2\sigma_0^2)$. But then according to Bernstein's theorem L_3 and L_4 are independent. The question therefore is, whether if L_3 and L_4 are independent, then X_i have the distributions which were given in previous. Again from Bernstein's theorem follows that $Y_1 \sim \mathrm{N}(b_1, \sigma_0^2)$, $Y_2 \sim \mathrm{N}(b_2, \sigma_0^2)$, but the distribution of X_i is possible to get only with some theory concerning characteristic functions.

Chapter 6

Theory

Definition: A characteristic function $f(t)$ is said to be *decomposable*, if it can be written in the form

$$f(t) = f_1(t)f_2(t),$$

where $f_1(t)$ and $f_2(t)$ are both characteristic functions of non-degenerate distributions. We say that $f_1(t)$ and $f_2(t)$ are *factors* of $f(t)$.

Definition: Let $z = t + iy, t, y \in \mathbb{R}, z \in \mathbb{C}$. A characteristic function is said to be *analytic characteristic function*, if there exists a function $A(z)$ of the complex variable z which is regular in the circle $|z| < \rho, \rho > 0$ and a constant r such that $A(t) = f(t)$ for $|t| < r$.

Remark: It is valid: If an analytic characteristic function is regular in some neighbourhood of the origin of the $\mathbb{C}$-plane, then it is also regular in a horizontal strip including the real line. Proof can be found in [4].

Definition: If the strip of regularity of $f(z)$ is the whole $\mathbb{C}$-plane, then we say that $f(z)$ is an *entire characteristic function*.

Definition: Let $f(z)$ be an entire characteristic function. We denote by

$$M(r, f) := \max_{|z| \leq r} |f(z)|.$$

Then the *order* of an entire characteristic function $f(z)$ is defined as

$$\rho = \limsup_{r \to \infty} \frac{\log \log M(r, f)}{\log r}.$$

Theorem 1: Let$f(z)$ be an analytic characteristic function, which is regular in the strip $-\alpha < Im(y) < \beta$; $\alpha, \beta > 0$. Then any factor $f_1(z)$ of $f(z)$ is also an analytic characteristic function which is regular, at least in the strip of regularity of $f(z)$.

Proof: Let

$$f(t) = f_1(t) \cdot f_2(t),$$

where $f_1(t), f_2(t)$ are two non-degenerate factors of $f(t)$, then the corresponding distribution functions $F(t), F_1(t), F_2(t)$ satisfy the relation

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y)\, dF_2(y) = \int_{-\infty}^{\infty} F_2(x-y)\, dF_1(y)$$

according to convolution theorem. Then for $A, B, \alpha_1, \alpha_2 \in \mathbb{R}$, $A, B > 0$, $\alpha_2 > \alpha_1$ we have

$$F(\alpha_1) = \int_{-\infty}^{\infty} F_1(\alpha_1 - y)\, dF_2(y), \qquad F(\alpha_2) = \int_{-\infty}^{\infty} F_1(\alpha_2 - y)\, dF_2(y),$$

$$F(\alpha_2) > F(\alpha_1) \qquad \Longrightarrow$$

$$F(\alpha_2) - F(\alpha_1) \geq \int_{-A}^{B} \left[F_1(\alpha_2 - y) - F_1(\alpha_1 - y)\right] dF_2(y). \tag{6.1}$$

From the theory concerning analytic characteristic functions we know, that for $v \in \mathbb{R}$, $-\alpha < v < \beta$, the integral

$$\int_{-\infty}^{\infty} e^{vx} dF(x)$$

exists and is finite. Clearly for $a, b \in \mathbb{R}$, $a < b$ is valid

$$\int_{-\infty}^{\infty} e^{vx} dF(x) \geq \int_{a}^{b} e^{vx} dF(x).$$

Now we will consider a sequence of subdivisions of the interval $[a, b]$:

$$x_j^{(n)} = a + \frac{b-a}{2^n}, (j-1) \qquad j = 1, 2, ..., 2^n + 1, \qquad n = 1, 2, ...,$$

so that we can representate the last integral by Darboux sums:

$$\int_{a}^{b} e^{vx} dF(x) = \lim_{n \to \infty} \sum_{j=1}^{2^n} \exp\left(v x_{j+1}^{(n)}\right) \left[F\left(x_{j+1}^{(n)}\right) - F\left(x_j^{(n)}\right)\right].$$

We denote by

$$\begin{aligned} h_{j,n}(y;v) &= \exp\left(vx_j^{(n)}\right)\left[F_1\left(x_{j+1}^{(n)}-y\right)-F_1\left(x_j^{(n)}-y\right)\right] \text{ for } v>0 \\ &= \exp\left(vx_{j+1}^{(n)}\right)\left[F_1\left(x_{j+1}^{(n)}-y\right)-F_1\left(x_j^{(n)}-y\right)\right] \text{ for } v<0 \end{aligned}$$

and by

$$g_n(y,v)=\sum_{j=1}^{2^n} h_{j,n}(y,v).$$

From (6.1) follows that

$$F\left(x_{j+1}^{(n)}\right)-F\left(x_j^{(n)}\right)\geq \int_{-A}^{B}\left[F_1\left(x_{j+1}^{(n)}-y\right)-F_1\left(x_j^{(n)}-y\right)\right]dF_2(y),$$

so that

$$\int_a^b e^{vx}dF(x)=\lim_{n\to\infty}\sum_{j=1}^{2^n}\exp\left(vx_{j+1}^{(n)}\right)\left[F\left(x_{j+1}^{(n)}\right)-F\left(x_j^{(n)}\right)\right]\geq \tag{6.2}$$

$$\geq \lim_{n\to\infty}\sum_{j=1}^{2^n}\exp\left(vx_{j+1}^{(n)}\right)\int_{-A}^{B}F_1\left(x_{j+1}^{(n)}-y\right)-F_1\left(x_j^{(n)}-y\right)dF_2(y)=$$

$$=\lim_{n\to\infty}\int_{-A}^{B}g_n(y,v)\,dF_2(y).$$

Now we return to the qualities of the division $\{x_j^{(n)}\}_{j=1}^{2^n}$. Clearly

$$x_{2j-1}^{(n+1)}<x_{2j}^{(n+1)}<x_{2j+1}^{(n+1)}$$

and also

$$x_{2j-1}^{(n+1)}=x_j^{(n)},$$

because $x_{2j-1}^{(n+1)}=a+\frac{b-a}{2^n}\cdot\frac{1}{2}\cdot(2j-2)=x_j^{(n)}$.
We note that $g_n(y,v)$ is a sum of 2^n terms, while $g_{n+1}(y,v)$ is a sum of twice as much terms, so we will compare terms

$$\begin{aligned} k &:= h_{j,n}(y;v)=\exp\left(vx_j^{(n)}\right)\left[F_1\left(x_{j+1}^{(n)}-y\right)-F_1\left(x_j^{(n)}-y\right)\right] \\ l &:= h_{2j,n+1}(y;v)=\exp\left(vx_{2j}^{(n+1)}\right)\left[F_1\left(x_{2j+1}^{(n+1)}-y\right)-F_1\left(x_{2j}^{(n+1)}-y\right)\right] \\ m &:= h_{2j-1,n+1}(y;v)=\exp\left(vx_{2j-1}^{(n+1)}\right)\left[F_1\left(x_{2j}^{(n+1)}-y\right)-F_1\left(x_{2j-1}^{(n+1)}-y\right)\right] \end{aligned}$$

for $v > 0$, because for $v < 0$ are the relations identical. Because $x_{2j-1}^{(n+1)} = = x_j^{(n)}$, we have $m - k = \exp\left(vx_{2j-1}^{(n+1)}\right)\left[F_1\left(x_{2j}^{(n+1)} - y\right) - F_1\left(x_{j+1}^{(n)} - y\right)\right]$ and therefore $l+m-k = \exp\left(vx_{2j}^{(n+1)}\right)\left[F_1\left(x_{2j+1}^{(n+1)} - y\right) - F_1\left(x_{2j}^{(n+1)} - y\right)\right] + + \exp\left(vx_{2j-1}^{(n+1)}\right)\left[F_1\left(x_{2j}^{(n+1)} - y\right) - F_1\left(x_{j+1}^{(n)} - y\right)\right] \geq \exp\left(vx_{2j-1}^{(n+1)}\right) \cdot \left[F_1\left(x_{2j+1}^{(n+1)} - y\right) - F_1\left(x_{j+1}^{(n)} - y\right)\right] = 0$. This implies that

$$h_{j,n}(y;v) \leq h_{2j,n+1}(y;v) + h_{2j-1,n+1}(y;v)$$

and therefore also $g_n(y,v) \leq g_{n+1}(y,v)$. So the sequence $\{g_n(y,v)\}_{n=1}^{\infty}$ is non-decreasing, it's terms are non-negative, because they are sums of non-negative numbers, therefore according to monotone convergence theorem (see [2]) we can write

$$\lim_{n\to\infty}\int_{-A}^{B} g_n(y,v)\,dF_2(y) = \int_{-A}^{B}\lim_{n\to\infty} g_n(y,v)\,dF_2(y). \tag{6.3}$$

The functions $g_n(y,v)$ are Darboux sums from definition and

$$\lim_{n\to\infty} g_n(y,v) = \int_{a-y}^{b-y} e^{v(y+z)}dF_1(z).$$

If we take this equation together with (6.2) and (6.3), we get

$$\int_a^b e^{vx}dF(x) \geq \lim_{n\to\infty}\int_{-A}^{B} g_n(y,v)\,dF_2(y) = \int_{-A}^{B}\lim_{n\to\infty} g_n(y,v)\,dF_2(y) =$$

$$= \int_{-A}^{B}\left(\int_{a-y}^{b-y} e^{v(y+z)}dF_1(z)\right)dF_2(y) = \int_{-A}^{B} e^{vy}\left[\int_{a-y}^{b-y} e^{vz}dF_1(z)\right]dF_2(y).$$

Because $y \in (-A,B)$, then $b-y \geq b-B$ and $a-y \leq a+A$ and thence

$$\int_{a-y}^{b-y} e^{vz}dF_1(z) \geq \int_{a+A}^{b-B} e^{vz}dF_1(z).$$

But this integral does not depend on y, so we can write

$$\infty > \int_{-\infty}^{\infty} e^{vx}dF(x) \geq \int_a^b e^{vx}dF(x) \geq \left[\int_{a+A}^{b-B} e^{vz}dF_1(z)\right]\left[\int_{-A}^{B} e^{vy}dF_2(y)\right].$$

This inequality is valid independently on a, b, A, B, so that

$$\int_{-\infty}^{\infty} e^{vx}dF(x) \geq \left[\int_{-\infty}^{\infty} e^{vz}dF_1(z)\right]\left[\int_{-\infty}^{\infty} e^{vy}dF_2(y)\right],$$

where both integrals on the right side exist and are finite. Here v is a real number, $-\alpha < v < \beta$, so the mentioned integrals exist and are finite for all such v. Then for $z \in \mathbb{C}$, $-\alpha < Im(z) < \beta$ the integrals

$$f_1(z) = \int_{-\infty}^{\infty} e^{izx} dF_1(x), \qquad f_2(z) = \int_{-\infty}^{\infty} e^{izx} dF_2(x)$$

exist and are finite, so we can see that $f_1(z)$ and $f_2(z)$ are analytic characteristic functions, whose strip of regularity is at least the strip of regularity of $f(z)$. Moreover, the equation $f(t) = f_1(t) f_2(t)$, which holds for real t, is also valid in the entire strip of regularity of $f(z)$.

Theorem 2: Let $f(z)$ be a decomposable analytic characteristic function with strip of regularity $-\alpha < Im(z) < \beta$. If $f_1(t)$ is a factor of $f(z)$, then there exist positive constants C and a such that

$$f_1(-iv) \leq Ce^{a|v|} f(-iv)$$

for all v satisfying $-\alpha < v < \beta$.

Proof: From the previous theorem we get

$$\begin{aligned} f(-iv) &= \int_{-\infty}^{\infty} e^{vx} dF(x) \geq \int_{-\infty}^{\infty} e^{vz} dF_1(z) \int_{-\infty}^{\infty} e^{vy} dF_2(y) = \\ &= f_1(-iv) \int_{-\infty}^{\infty} e^{vy} dF_2(y). \end{aligned}$$

We choose two real numbers a_1,a_2 such that $0 < F_2(a_1)$, $1 > F_2(a_2)$. Then

$$\int_{-\infty}^{\infty} e^{vy} dF_2(y) \geq \int_{a_2}^{\infty} e^{vy} dF_2(y) \geq \int_{a_2}^{\infty} e^{a_2 v} dF_2(y) = e^{a_2 v} [1 - F_2(a_2)]$$

for $v > 0$

$$\int_{-\infty}^{\infty} e^{vy} dF_2(y) \geq \int_{-\infty}^{a_1} e^{vy} dF_2(y) \geq \int_{-\infty}^{a_1} e^{a_1 v} dF_2(y) = e^{a_1 v} [F_2(a_1)] \text{ for } v < 0$$

Let $a = \max[|a_1|, |a_2|]$. Then surely $e^{-a|v|} \leq e^{a_1 v}$, $e^{-a|v|} \leq e^{a_2 v}$ $\forall v \in (-\alpha, \beta)$. But if we denote by $C^{-1} = \min[1 - F_2(a_2), F_2(a_1)]$, then

$$\int_{-\infty}^{\infty} e^{vy} dF_2(y) \geq C^{-1} e^{-a|v|} \text{ and } f(-iv) \geq C^{-1} e^{-a|v|} f_1(-iv),$$

from which immediately follows the statement given in Theorem 2.

Lemma: Let $f(z)$ be an analytic characteristic function. Then $|f(z)|$ attains it's maximum along any horizontal line contained in the interior of it's strip of regularity on the imaginary axis. The derivatives $f^{(2k)}(z)$, $k \in \mathbb{Z}$ of $f(z)$ have the same property.

Proof:

$$f(z) = \int_{-\infty}^{\infty} e^{izx} dF(x), \qquad -\alpha < Im(z) < \beta$$

Since $f(z)$ is an analytic characteristic function, then all moments of the corresponding distribution exist. Then we see that the integral

$$\int_{-\infty}^{\infty} i^r x^r e^{izx} dF(x), \qquad r \in \mathbb{Z}$$

has an integrable majorant and we can therefore derivate under the integral sign (for proof see [3])

$$f^{(r)}(z) = i^r \int_{-\infty}^{\infty} x^r e^{izx} dF(x), \qquad z = a + iy,\ a \in \mathbb{R},\ y \in (-\alpha, \beta)$$

$$|f^{(r)}(a+iy)| \le \int_{-\infty}^{\infty} |x^r||e^{ix(a+iy)}| dF(x) = \int_{-\infty}^{\infty} |x^r| e^{-xy} dF(x)$$

so if we take $r = 2k$, $k = 0, 1, 2, \dots$, we get

$$|f^{(2k)}(a+iy)| \le \int_{-\infty}^{\infty} x^{2k} e^{-xy} dF(x) = |f^{(2k)}(iy)|$$

so that

$$\max_{-\infty < a < \infty} |f^{(2k)}(a+iy)| = |f^{(2k)}(iy)|.$$

Remark: From this lemma immediately follows, that

$$M(r, f) = \max\left[f(ir), f(-ir)\right].$$

Theorem 3: Every factor of an entire characteristic function is an entire characteristic function. The order of this factor cannot exceed the order of the original characteristic function.

Proof: Theorem 1 says that every factor of an analytic characteristic function is an analytic characteristic function, which is regular at least in the strip of regularity of the original characteristic function. Therefore clearly any factor of an entire characteristic function is an entire characteristic function. From the previous remark we have $M(r,f) = \max[f(ir), f(-ir)]$ and from Theorem 2 $f_1(-iv) \leq Ce^{a|v|}f(-iv)$, where f_1 is a factor of f. Therefore evidently

$$M(r,f_1) \leq M(r,f)\,Ce^{ar}$$

and we see that the order of f_1 cannot exceed the order of f.

Theorem 4 (Cramer): The characteristic function

$$f(t) = \exp\{i\mu t - \frac{t^2\sigma^2}{2}\}$$

of normal distribution has only normal factors. Moreover, if

$$f(t) = f_1(t)\, f_2(t), \qquad \text{where} \qquad f_j(t) = \exp\{i\mu_j t - \frac{t^2\sigma_j^2}{2}\}, \qquad j = 1,2$$

then $\mu_1 + \mu_2 = \mu$ and $\sigma_1^2 + \sigma_2^2 = \sigma^2$.

Proof: The function $f(t)$ is an entire characteristic function that does not attend 0, so according to Theorem 3 the factors of $f(t)$ have the same property. Therefore $f_1(t)$ has the form

$$f_1(t) = \exp\{g_1(t)\}$$

and it follows from the Marcinkiewicz theorem (proof can be found in [4]) that $g_1(t)$ is a polynomial of degree not exceeding 2. So let

$$g_1(t) = a_0 + a_1 t + a_2 t^2.$$

Since $f_1(0) = 1$, we see that $a_0 = 0$. Because $f_1(t)$ must satisfy the relation $\overline{f_1(t)} = f_1(-t)$, the same must be valid for $g_1(t)$, so $a_1 = i\mu_1$ and $a_2 \in \mathbb{R}$. Further,

$$1 \geq |f_1(t) = \exp\{a_2 t^2\},$$

therefore $a_2 \leq 0$ and we set $a_2 = -\frac{\sigma_1^2}{2}$. Together we get

$$f_1(t) = \exp\{i\mu_1 t - \frac{t^2\sigma_1^2}{2}\}.$$

The same must be valid for $f_2(t)$ and it is easy to see that $\mu_1 + \mu_2 = \mu$ and $\sigma_1^2 + \sigma_2^2 = \sigma^2$, because

$$\exp\{i\mu t - \frac{t^2\sigma^2}{2}\} = \exp\{i\mu_1 t - \frac{t^2\sigma_1^2}{2}\} \cdot \exp\{i\mu_2 t - \frac{t^2\sigma_2^2}{2}\}.$$

Chapter 7

Modification of Bernstein's and Polya's theorem - continuation

From the Cramer's theorem we can see that $X_1, ..., X_4$ are normally distributed with $X_1 \sim \mathrm{N}(a_1, \sigma_1^2)$, $X_2 \sim \mathrm{N}(a_2, \sigma_2^2)$, $X_3 \sim \mathrm{N}(a_3, \sigma_3^2)$, $X_4 \sim \mathrm{N}(a_4, \sigma_4^2)$, where $\sigma_1^2 + \sigma_2^2 = \sigma_0^2 = \sigma_3^2 + \sigma_4^2$, $a_1 + a_2 = b_1$, $a_3 + a_4 = b_2$, and the modification of Bernstein's theorem is therefore proved.

We also note that with Cramer's theorem we can prove Polya's theorem in more general way: if we write again the equation

$$f(t) = f^2\left(\frac{t}{\sqrt{2}}\right) \qquad \text{or} \qquad f(-t) = f^2\left(\frac{-t}{\sqrt{2}}\right),$$

then by multiplying these two equations we get

$$|f(t)|^2 = |f\left(\frac{t}{\sqrt{2}}\right) \cdot f\left(\frac{-t}{\sqrt{2}}\right)|^2.$$

But now we can set $g\left(\frac{t}{\sqrt{2}}\right) := f\left(\frac{t}{\sqrt{2}}\right) \cdot f\left(\frac{-t}{\sqrt{2}}\right) = g\left(\frac{t}{\sqrt{2}}\right)$, so that $g(t)$ is a symmetric characteristic function, for which is valid

$$g(t) = |g\left(\frac{t}{\sqrt{2}}\right)|^2.$$

We see that $g(t)$ is real and from the yet proved version of Polya's theorem $g(t)$ is a characteristic function of $N(0, \sigma^2)$. But since $g(t) = f(t) \cdot f(-t)$, then from the Cramer's theorem follows that $f(t)$ is a characteristic function of $N\left(0, \frac{\sigma^2}{2}\right)$. So the presumption for symmetry of the characteristic function in Polya's theorem is not necessary.

Chapter 8

Applications

At the end let us see some uses of Polya's theorem:

1. Polya's theorem can be understood as some equivalent of central limit theorem: Let us have a sequence of iid random variables $\{X_i\}$, we denote
$$S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$$
and presume that there exists a random variable Y such, that $S_n \xrightarrow{d} Y$ for $n \to \infty$.Then also $S_{2n} \xrightarrow{d} Y$ for $n \to \infty$ and we can write
$$S_{2n} = \frac{1}{\sqrt{2n}} \sum_{i=1}^{2n} X_i = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i + \frac{1}{\sqrt{n}} \sum_{i=n+1}^{2n} X_i \right). \tag{8.1}$$
We know that
$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \xrightarrow{d} Y \text{ for } n \to \infty, \qquad \frac{1}{\sqrt{n}} \sum_{i=n+1}^{2n} X_i \xrightarrow{d} Y' \text{ for } n \to \infty,$$
where $Y \stackrel{d}{=} Y'$. Then we have (after applying $\lim n \to \infty$ on both sides of (8.1))
$$Y \stackrel{d}{=} \frac{1}{\sqrt{2}} \left(Y + Y' \right),$$
but according to Polya's theorem Y must have normal distribution with $\mathrm{N}(0, \sigma^2)$.

2. Polya's theorem can be also used when we want to model the movement of the Brown's particle in some liquid. If we set it's movement's trajectory to the $\mathbb{R}^2$-plane, then on the axes we can observe the values of random variables Y_1 and Y_2, which we can presume are iid with the distribution $\mathrm{N}(0, \sigma^2)$. If we now shift the axes by an angle of $\frac{\pi}{2}$ in the anti-clockwise sense, the random variable Y_1 will be transformed to random variable $\frac{Y_1+Y_2}{\sqrt{2}}$, which has according to Polya's theorem the same distribution as Y_1.

Chapter 9

Conclusion

This thesis shows us the importance and usefulness of the concept of characteristic functions. They are not only a useful tool for proving characterization theorems, but we can also understand the connection between particular characterization theorems better due to characteristic functions, because they enable us to re-write the characterization theorems in the form of functional equations and thus give us the possibility to use the huge arsenal of functional analysis for getting significant results in probability theory and mathematical statistics.

Bibliography

[1] Lachout P.: *Teorie pravdepodobnosti*, Karolinum, Praha, 2004.

[2] Loeve M.: *Probability theory*, D. Van Nostrand, New York, 1963.

[3] Cramer H.: *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1964.

[4] Lukacs E.: *Characteristic functions*, Griffin, London, 1970