Chapter 2

Weak Convergence

Characteristic Functions

If we have a probability distribution α on the line then its characteristic function is defined by

$$\phi(t) = \int \exp[\, i\, t\, x\,]\, d\alpha \tag{2.1}$$

The above definition makes sense. We write the integrand e^{itx} as $\cos tx + i\, \sin tx$ and integrate each part to see that

$$|\phi(t)| \leq 1$$

for all real t.

Exercise 2.1. Calculate the characteristic functions for the following distributions:

1. α is the degenerate distribution δ_a with probability one at the point a.

2. α is the binomial distribution masses

   $$p_k = Prob[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

   for $0 \leq k \leq n$.

Theorem 2.1. *The characteristic function $\phi(t)$ of any probability distribution is a uniformly continuous function of t that is positive definite i.e. for any n complex numbers $\xi_1, \cdots, \xi_n$ and real nubers $t_1, \cdots, t_n$*

$$\sum_{i,j=1}^{n} \phi(t_i - t_j)\, \xi_i\, \bar{\xi}_j \geq 0$$

Proof: Let us note that

$$\begin{aligned}\sum_{i,j=1}^{n} \phi(t_i - t_j)\, \xi_i\, \bar{\xi}_j \quad &= \textstyle\sum_{i,j=1}^{n} \xi_i\, \bar{\xi}_j \int \exp[\, i(t_i - t_j)x]\, d\alpha \\ &= \textstyle\int |\sum_{j=1}^{n} \xi_j \exp[i\, t_j\, x]|^2\, d\alpha \geq 0\end{aligned}$$

To prove uniform continuity we see that

$$|\phi(t) - \phi(s)| \leq \int |\exp[\, i(t-s)\, x\,] - 1|\, dP$$

which tends to 0 by the bounded convergence theorem if $|t - s| \to 0$.

The characteristic function of course carries some information about the distribution α. In particular if $\int |x|\, d\alpha < \infty$, then $\phi(\cdot)$ is continuously diffrentiable and $\phi'(0) = i \int x\, d\alpha$.

Exercise 2.2. Prove it!

Warning: The converse need not be true. $\phi(\cdot)$ can be continuously differentiable but $\int |x|\, dP$ could be ∞.

Exercise 2.3. Construct a counterexample along the followimg lines. Take a discrete distribution, symmetric around 0 with

$$\alpha\{n\} = \alpha\{-n\} = p(n) \simeq \frac{1}{n^2 \log n}$$

Then show that $\sum_n \frac{(1-\cos nt)}{n^2 \log n}$ is a continuously differentiable function of t.

Exercise 2.4. The story with higher moments $m_r = \int x^r\, d\alpha$ is similar. If any of them, say m_r exists, then $\phi(\cdot)$ is r times continuously differentiable and $\phi^{(r)}(0) = i^r m_r$. The converse is false for odd r, but true for even r.

The next question is how to recover the distribution function $F(x)$ from $\phi(t)$. If we go back to the Fourier inversion formula, we can 'guess' that

$$F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-i\,t\,x\,]\,\phi(t)\,dt$$

and therefore

$$\begin{aligned}
F(b) - F(a) &= \tfrac{1}{2\pi}\int_a^b dx \int_{-\infty}^{\infty} \exp[-i\,t\,x\,]\,\phi(t)\,dt \\
&= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t)\,dt \int_a^b \exp[-i\,t\,x\,]\,dx \\
&= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t)\tfrac{\exp[-\,i\,t\,b\,]-\exp[-\,i\,t\,a\,]}{-\,i\,t}\,dt \\
&= \lim_{T\to\infty} \tfrac{1}{2\pi}\int_{-T}^{T} \phi(t)\tfrac{\exp[-\,i\,t\,b\,]-\exp[-\,i\,t\,a\,]}{-\,i\,t}
\end{aligned}$$

We will in fact prove it if a and b are points of continuity of F. We compute the right hand side as

$$\lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{\exp[-\,i\,t\,b\,]-\exp[-\,i\,t\,a\,]}{-\,i\,t}\,dt \int \exp[\,i\,t\,x\,]\,d\alpha$$

$$\begin{aligned}
&= \lim_{T\to\infty} \tfrac{1}{2\pi}\int d\alpha \int_{-T}^{T} \tfrac{\exp[i\,t\,(x-b)\,]-\exp[i\,t\,(x-a)\,]}{-\,i\,t}\,dt \\
&= \lim_{T\to\infty} \tfrac{1}{2\pi}\int d\alpha \int_{-T}^{T} \tfrac{\sin t\,(x-a)-\sin t\,(x-b)}{t}\,dt \\
&= \tfrac{1}{2}\int [\operatorname{sign}(x-a) - \operatorname{sign}(x-b)]\,d\alpha \\
&= F(b) - F(a)
\end{aligned}$$

provided a and b are continuity points. We have applied the bounded convergence theorem to take the limit as $T \to \infty$.

As a consequence we conclude that the distribution function and hence α is determined uniquely by the characteristic function.

Weak Convergence

One of the basic ideas in Limit Theorems is the notion of **weak Convergence** of a sequence of probability distributions on the line **R**. Since the role of a probability measure is to assign probabilities to sets, we should expect that if two probability measures are to be close then they should assign for a given set, probabilities that are nearly equal. This suggets the definition

$$d(P_1, P_2) = \sup_{A \in \mathcal{B}} |P_1(A) - P_2(A)|$$

as the distance between two probability measures P_1 and P_2 on a measurable space $(\Omega, \mathcal{B})$. This is too strong. If we take P_1 and P_2 to be degenerate distributions with probability 1 concentrated at two points x_1 and x_2 on the line one can see that, as soon as $x_1 \neq x_2$, $d(P_1, P_2) = 1$, and the above metric is not sensitive to how close the two points x_1 and x_2 are. It only cares that they are unequal. The problem is not because of the supremum. We can take A to be an interval $[a, b]$ that includes x_1 but omits x_2 and $|P_1(A) - P_2(A)| = 1$. On the other hand if the end points of the interval are kept away from x_1 or x_2 the situation is not that bad. This leads to the following definition.

Definition 2.1. A sequence α_n of probabiity distributions on $\mathbf{R}$ is said to converge weakly to a probability distribution α if,

$$\lim_{n \to \infty} \alpha_n[I] = \alpha[I]$$

for any interval $I = [a, b]$ such that the single point sets a and b have probability 0 under α.

Remark. One says that a sequence X_n of random variables **converges in law** or in distribution to X if the distributions α_n of X_n converges weakly to α the distribution of X.

One can state this more succinctly in terms of the distribution functions $F_n(x)$ and $F(x)$ corrresponding to the measures α_n and α respectively.

Definition 2.2. A sequence α_n of probability measures on the real line $\mathbf{R}$ with distribution functions $F_n(x)$ is said to converge weakly to a limiting probability measure α weakly (in symbols $\alpha_n \Rightarrow \alpha$ or $F_n \Rightarrow F$) if

$$\lim_{n \to \infty} F_n(x) = F(x)$$

for every x that is a continuity point of F.

There are equivalent formulations interms of expectations and characteristic functions.

Theorem 2.2. *The following are equivalent.*

1. $\alpha_n \Rightarrow \alpha$ *or* $F_n \Rightarrow F$

2. *For every bounded continuous function* $f(x)$ *on* $\mathbf{R}$

$$\lim_{n\to\infty}\int_{\mathbf{R}} f(x)\,d\alpha_n = \int_{\mathbf{R}} f(x)\,d\alpha$$

3. *If* $\phi_n(t)$ *and* $\phi(t)$ *are respectively the characteristic functions of* α_n *and* α*, for every real* t*,*

$$\lim_{n\to\infty}\phi_n(t) = \phi(t)$$

Proof: We first prove $(a \Rightarrow b)$. Let $\epsilon > 0$ be arbitrary. Find continuity points a and b of F such that $a < b$, $F(a) \le \epsilon$ and $1 - F(b) \le \epsilon$. Since $F_n(a)$ and $F_n(b)$ converge to $F(a)$ and $F(b)$, for n large enough, $F_n(a) \le 2\epsilon$ and $1 - F_n(b) \le 2\epsilon$. Divide the interval $[a, b]$ into a finite number $N = N_\delta$ of small subintervals $I_j = (a_j, a_{j+1}], 1 \le j \le N$ with $a = a_1 < a_2 < \cdots < a_{N+1} = b$ such that all the end points $\{a_j\}$ are points of continuity of F and the oscillation of the continuous function f in each I_j is less than a preassigned number δ. Since the continuous function f is uniformly continuous in the closed bounded interval $[a, b]$ this is always possible for any given $\delta > 0$. Let $h(x) = \sum_{j=1}^{N}\chi_{I_j} f(a_j)$ be the simple function equal to $f(a_j)$ on I_j and 0 outside $\cup_j I_j = (a, b]$. We have $|f(x) - h(x)| \le \delta$ on $(a, b]$. If $f(x)$ is bounded by M, then

$$\Big|\int f(x)\,d\alpha_n - \sum_{j=1}^{N} f(a_j)[F_n(a_{j+1}) - F_n(a_j)]\Big| \le \delta + 4M\epsilon$$

and

$$\Big|\int f(x)\,d\alpha - \sum_{j=1}^{N} f(a_j)[F(a_{j+1}) - F(a_j)]\Big| \le \delta + 2M\epsilon$$

Since $\lim_{n\to\infty} F_n(a_j) = F(a_j)$ for every $1 \le j \le N$, we conclude that

$$\limsup_{n\to\infty}\Big|\int f(x)\,d\alpha_n - \int f(x)\,d\alpha\Big| \le 2\delta + 6M\epsilon$$

Since ϵ and δ are arbitrary small numbers we are done. $(b \Rightarrow c)$ is trivial. We can make the choice of $f(x) = \exp[i\,t\,x\,] = \cos tx + i \sin tx$ which is a bounded continuous function. $(c \Rightarrow a)$ is the hardest. It is carried out in several steps. Actually we will prove a stronger version as a separate theorem.

Theorem 2.3. For each $n \geq 1$, let $\phi_n(t)$ be the characteristic function of a probability distribution α_n. Assume that $\lim_{n\to\infty} \phi_n(t) = \phi(t)$ exists for each t and $\phi(t)$ is continuous at $t = 0$. Then $\phi(t)$ is the characteristic function of some probability distribution α and $\alpha_n \Rightarrow \alpha$.

Proof:

Step 1. Let $r_1, r_2, \cdots$ be an enumeration of the rational numbers. For each j consider the sequence $\{F_n(r_j) : n \geq 1\}$ where F_n is the distribution function corresponding to $\phi_n(\cdot)$. It is a sequence bounded by 1 and we can extract a subsequence that converges. By the diagonalization process we can choose a subseqence $G_k = F_{n_k}$ such that

$$\lim_{n\to\infty} G_k(r) = b_r$$

exists for every rational number r. From the monotonicity of F_n in x we conclude that if $r_1 < r_2$, then $b_{r_1} \leq b_{r_2}$.

Step 2. From the skeleton b_r we reconstruct a right continuous monotone function $G(x)$. We defiine

$$G(x) = \inf_{r>x} b_r$$

Clearly G is nondecreasing. If $x_n \downarrow x$, any $r > x$ satisfies $r > x_n$ for sufficiently large n. This allows us to conclude that $G(x) = \inf_n G(x_n)$ for any sequence $x_n \downarrow x$, proving that $G(x)$ is right continuous.

Step 3. Next we show that at any continuity point x of G

$$\lim_{n\to\infty} G_n(x) = G(x)$$

Let $r > x$ be a rational number. Then $G_n(x) \leq G_n(r)$ and $G_n(r) \to b_r$ as $n \to \infty$. Hence

$$\limsup_{n\to\infty} G_n(x) \leq b_r$$

This is true for every rational $r > x$, and therefore

$$\limsup_{n\to\infty} G_n(x) \le G(x)$$

Suppose now that we have $y < x$. Find a rational r such that $y < r < x$.

$$\liminf_{n\to\infty} G_n(x) \ge \liminf_{n\to\infty} G_n(r) = b_r \ge G(y)$$

As this is true for every $y < x$,

$$\liminf_{n\to\infty} G_n(x) \ge \sup_{y<x} G(y) = G(x-0) = G(x)$$

the last step being a consequence of the assumption that x is a point of continuity of G i.e. $G(x-0) = G(x)$.

Warning. This does not mean that G is necessarily a distribution function. Consider $F_n(x) = 0$ for $x < n$ and 1 for $x \ge n$. In this case $\lim_{n\to\infty} F_n(x) = G(x)$ exists and $G(x) \equiv 0$, which is not a distribution function.

Step 4. We will use the continuity at $t = 0$, of $\phi(t)$, to show that G is a distribution function. If $\phi(t)$ is the characteristic function of α

$$\begin{aligned}
\tfrac{1}{2T}\textstyle\int_{-T}^{T}\phi(t)\,dt &= \textstyle\int\left[\tfrac{1}{2T}\int \exp[i\,t\,x\,]\,dt\right]\,d\alpha \\
&= \textstyle\int \tfrac{\sin Tx}{Tx}\,d\alpha \\
&\le \textstyle\int \left|\tfrac{\sin Tx}{Tx}\right|\,d\alpha \\
&= \textstyle\int_{|x|<\ell}\left|\tfrac{\sin Tx}{Tx}\right|\,d\alpha + \int_{|x|\ge\ell}\left|\tfrac{\sin Tx}{Tx}\right|\,d\alpha \\
&\le \alpha[|x|<\ell] + \tfrac{1}{T\ell}\,\alpha[|x|\ge\ell]
\end{aligned}$$

We have used Fubini's theorem in the first line and the bounds $|\sin x| \le |x|$ and $|\sin x| \le 1$ in the last line. We can rewrite this as

$$\begin{aligned}
1 - \tfrac{1}{2T}\textstyle\int_{-T}^{T}\phi(t)\,dt &\ge 1 - \alpha[|x|<\ell] - \tfrac{1}{T\ell}\,\alpha[|x|\ge\ell] \\
&= (1 - \tfrac{1}{T\ell})\alpha[|x|\ge\ell] \\
&\ge (1 - \tfrac{1}{T\ell})[1 - F(\ell) + F(-\ell)]
\end{aligned}$$

Finally, if we pick $\ell = \frac{2}{T}$,

$$[1 - F(\frac{2}{T}) + F(-\frac{2}{T})] \leq 2\Big[1 - \frac{1}{2T}\int_{-T}^{T}\phi(t)\,dt\Big]$$

Since this inequality is valid for any distribution function F and its characteristic function ϕ, we conclude that, for every $n \geq 1$,

$$[1 - F_{n_k}(\frac{2}{T}) + F_{n_k}(-\frac{2}{T})] \leq 2\Big[1 - \frac{1}{2T}\int_{-T}^{T}\phi_{n_k}(t)\,dt\Big].$$

We can pick T such that $\pm\frac{2}{T}$ are continuity points of G. If we now pass to the limit and use the bounded convergence theorem, we obtain

$$[1 - G(\frac{2}{T}) + G(-\frac{2}{T})] \leq 2\Big[1 - \frac{1}{2T}\int_{-T}^{T}\phi(t)\,dt\Big]$$

Since $\phi(0) = 1$ and ϕ is continuous at $t = 0$, by letting $T \to 0$ in such a way that $\pm\frac{2}{T}$ are continuity points of G, we conclude that

$$1 - G(\infty) + G(-\infty) = 0$$

or G is indeed a distribution function

Step 4. We now complete the rest of the proof, i.e. show that $\alpha_n \Rightarrow \alpha$. We have $G_k = F_{n_k} \Rightarrow G$ as well as $\psi_k = \phi_{n_k} \to \phi$. Therefore G must equal F which has ϕ for its characteristic function. Since the argument works for any subsequence of F_n, every subsequence of F_n will have a further subsequence that converges weakly to the same limit F uniquely determined as the distribution function whose characteristic function is $\phi(\cdot)$. Consequently $F_n \Rightarrow F$ or $\alpha_n \Rightarrow \alpha$.

Definition 2.3. A subset $\mathcal{A}$ of probability distributions on $\mathbf{R}$ is said to be totally bounded if, given any sequence α_n from $\mathcal{A}$, there is subsequence that converges weakly to some limiting probability distribution α.

Theorem 2.4. *In order that a family $\mathcal{A}$ of probability distributions to be totally bounded it is necessary and sufficient that either of the following equivalent conditions hold.*

a)

$$\lim_{\ell\to\infty}\sup_{\alpha\in\mathcal{A}}\alpha[x : |x|\geq \ell] = 0$$

b)

$$\lim_{h\to 0}\sup_{\alpha\in\mathcal{A}}\sup_{|t|\leq h}|1-\phi_\alpha(t)| = 0$$

Here $\phi_\alpha(t)$ is the characteristic function of α.

Condition a) is often called the uniform tightness property.

Proof. The proof is already contained in the details of the proof of the earlier theorem. We can always choose a subsequence such that the distribution functions converge at rationals and try to reconstruct the limiting distribution function from the limits at rationals. The crucial step is to prove that the limit is a distribution function. Either of the two conditions a) or b) will guarantee this. If a) is violated it is straight forward to pick a sequence from $\mathcal{A}$ for which the distribution functions have a limit which is not a distribution function. Then $\mathcal{A}$ cannot be totally bounded. a) is therefore necessary. That a)$\Rightarrow$ b), is a consequence of the estimate

$$\begin{aligned}
|1-\phi(t)| &\leq \textstyle\int |\exp[\,i\,t\,x\,]-1|\,d\alpha \\
&= \textstyle\int_{|x|\leq\ell}|\exp[\,i\,t\,x\,]-1|\,d\alpha + \int_{|x|>\ell}|\exp[\,i\,t\,x\,]-1|\,d\alpha \\
&\leq |t|\ell + 2\alpha[x : |x|>\ell]
\end{aligned}$$

The regularity of $\phi(t)$ at $t=0$ is related to the decay rate of the tail probabilities.

Exercise 2.5. Compute $\int |x|^p\,d\alpha$ in terms of the characteristic function $\phi(t)$ for p in the range $0<p<2$.

Hint: Look at the formula

$$\int_{-\infty}^{\infty}\frac{1-\cos tx}{|t|^{p+1}}dt = C_q|x|^p$$

and use Fubini's theorem.

Bochner's Theorem

We are now ready to prove the converse of Theorem 2.1 due to Bochner.

Theorem 2.5. **(Bochner)**. *If $\phi(t)$ is a positive definite function which is continuous at $t = 0$ and is normalized so that $\phi(0) = 1$, then ϕ is the characteristic function of some probability ditribution on* $\mathbf{R}$.

Proof: The proof depends on constructing approximations $\phi_n(t)$ which are in fact characteristic functions and satisfy $\phi_n(t) \to \phi(t)$ as $n \to \infty$. Then we can apply the preceeding theorem and the probability measures corresponding to ϕ_n will have a weak limit which will have ϕ for its characteristic function.

Step 1. Let us establish a few elementary properties of positive definite functions.

1) If $\phi(t)$ is a positive definite function so is $\phi(t) exp[\, i\, t\, a\,]$ for any real a. The proof is elementary and requires just direct verification.

2) If $\phi_j(t)$ are positive definite for each j then so is any linear combination $\phi(t) = \sum_j w_j \phi_j(t)$ with nonnegative weights w_j. If each $\phi_j(t)$ is normalized with $\phi_j(0) = 1$ and $\sum_j w_j = 1$, then of of course $\phi(0) = 1$ as well.

3) If ϕ is positive definite then ϕ satisfies $\phi(0) \geq 0$, $\phi(-t) = \overline{\phi(t)}$ and $|\phi(t)| \leq \phi(0)$ for all t.

We use the fact that the matrix $\{\phi(t_i - t_j) : 1 \leq i, j \leq n\}$ is Hermitian positive definite for any n real numbers $t_1, \cdots, t_n$. The first assertion follows from the the positivity of $\phi(0)|z|^2$, the second is a consequence of the Hermitian property and if we take $n = 2$ with $t_1 = t$ and $t_2 = 0$ as a consequence of the positive definiteness of the 2×2 matrix we get $|\phi(t)|^2 \leq |\phi(0)|^2$

4) For any s, t we have $|\phi(t) - \phi(s)|^2 \leq 4\phi(0)|\phi(0) - \phi(t - s)|$

We use the positive definiteness of the 3×3 matrix

$$\begin{matrix} 1 & \phi(t-s) & \phi(t) \\ \overline{\phi(t-s)} & 1 & \phi(s) \\ \overline{\phi(t)} & \overline{\phi(s)} & 1 \end{matrix}$$

which is $\{\phi(t_i - t_j)\}$ with $t_1 = t, t_2 = s$ and $t_3 = 0$. In particular the determinant has to be nonnegative.

$$\begin{aligned}
0 &\leq 1+\phi(s)\phi(t-s)\overline{\phi(t)}+\overline{\phi(s)\phi(t-s)}\phi(t)-|\phi(s)|^2 \\
&\quad -|\phi(t)|^2-|\phi(t-s)|^2 \\
&= 1-|\phi(s)-\phi(t)|^2-|\phi(t-s)|^2-\phi(t)\overline{\phi(s)}(1-\phi(t-s)) \\
&\quad -\overline{\phi(t)}\phi(s)(1-\overline{\phi(t-s)}) \\
&\leq 1-|\phi(s)-\phi(t)|^2-|\phi(t-s)|^2+2|1-\phi(t-s)|
\end{aligned}$$

Or

$$\begin{aligned}
|\phi(s)-\phi(t)|^2 &\leq 1-|\phi(s-t)|^2+2|1-\phi(t-s)| \\
&\leq 4|1-\phi(s-t)|
\end{aligned}$$

5) It now follows from 4) that if a positive definite function is continuous at $t=0$, it is continuous everywhere.

Step 2. First we show that if $\phi(t)$ is a positive definite function which is continuous on $\mathbf{R}$ and is absolutely integrable, then

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp[-\,i\,t\,x\,]\phi(t)\,dt\geq 0$$

is a continuous function and

$$\int_{-\infty}^{\infty}f(x)dx=1.$$

Moreover the function

$$F(x)=\int_{-\infty}^{x}f(y)\,dy$$

defines a distribution function with characteristic function

$$\phi(t)=\int_{-\infty}^{\infty}\exp[i\,t\,x\,]f(x)\,dx \tag{2.2}$$

If ϕ is integrable on $(-\infty,\infty)$, then $f(x)$ is clearly bounded and continuous. To see that it is nonnegative we write

$$\begin{aligned} f(x) &= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} (1 - \frac{|t|}{T}) e^{-\,i\,t\,x} \phi(t)\, dt &(2.3)\\ &= \lim_{T\to\infty} \frac{1}{2\pi T} \int_0^T \int_0^T e^{-\,i\,(t-s)\,x} \phi(t-s)\, dt\, ds &(2.4)\\ &= \lim_{T\to\infty} \frac{1}{2\pi T} \int_0^T \int_0^T e^{-\,i\,t\,x}\, e^{i\,s\,x} \phi(t-s)\, dt\, ds &(2.5)\\ &\geq 0 \end{aligned}$$

We can use the dominated convergence theorem to prove 2.3, a change of variables to show 2.4 and finally a Riemann sum approximation to the integral and the positive definiteness of ϕ to show that 2.5 is nonnegative. It reamains to show the relation 2.2. Let us define

$$f_\sigma(x) = f(x) \exp[\, -\frac{\sigma^2 x^2}{2}]$$

and calculate

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\,t\,x} f_\sigma(x)\, dx &= \int_{-\infty}^{\infty} e^{i\,t\,x} f(x) \exp[\, -\tfrac{\sigma^2 x^2}{2}]\, dx \\ &= \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\,t\,x} \phi(s) e^{i\,s\,x} \exp[\, -\tfrac{\sigma^2 x^2}{2}]\, ds\, dx \\ &= \int_{-\infty}^{\infty} \phi(s) \tfrac{1}{\sqrt{2\pi}\sigma} \exp[\, -\tfrac{(t-s)^2}{2\sigma^2}]\, ds \end{aligned}$$

by Fubini's thorem. It is now clear that

$$\int_{-\infty}^{\infty} f_\sigma(x)\, dx = \int_{-\infty}^{\infty} \phi(s) \frac{1}{\sqrt{2\pi}\sigma} \exp[\, -\frac{s^2}{2\sigma^2}]\, ds \leq 1$$

and as a consequence of the monotone convergence theorem, by letting $\sigma \to 0$, we have established our claim.

Step 3. If $\phi(t)$ is a positive definite function which is continuous, so is $\phi(t) \exp[\, i\,t\,y\,]$ for every y and for $\sigma > 0$, as well as the convex combination

$$\begin{aligned} \phi_\sigma(t) &= \int_{-\infty}^{\infty} \phi(t) \exp[\, i\,t\,y\,] \tfrac{1}{\sqrt{2\pi}\sigma} \exp[\, -\tfrac{y^2}{2\sigma^2}]\, dy \\ &= \phi(t) \exp[\, -\tfrac{\sigma^2 t^2}{2}] \end{aligned}$$

The previous step is applicable to $\phi_\sigma(t)$ which is clearly integrable on $\mathbf{R}$ and by letting $\sigma \to 0$ we conclude by Theorem 2.3. that ϕ is a characteristic function as well.

Theorem 2.6. *Let $\alpha_n \Rightarrow \alpha$ on $\mathbf{R}$ and C be a closed subset of $\mathbf{R}$. Then*

$$\limsup_{n\to\infty} \alpha_n(C) \leq \alpha(C).$$

If $A \subset \mathbf{R}$ is a continuity set of α i.e. $\alpha(\partial A) = \alpha(\bar{A} - A^o) = 0$, then

$$\lim_{n\to\infty} \alpha_n(A) = \alpha(A)$$

Proof: The function $d(x, C) = \inf_{y\in C} |x-y|$ is a continuous and equals 0 precisely on C.

$$f(x) = \frac{1}{1 + d(x, C)}$$

is a continuous function bounded by 1, that is equal to 1 precisely on C and

$$f_k(x) = [f(x)]^k \downarrow \chi_C(x)$$

as $k \to \infty$. From

$$\lim_{n\to\infty} \int f_k(x)\, d\alpha_n = \int f_k(x)\ d\alpha$$

for each $k \geq 1$, we deduce that

$$\limsup_{n\to\infty} \alpha_n(C) \leq \alpha(C)$$

Taking complements we conclude that for any open set $G \subset \mathbf{R}$

$$\liminf_{n\to\infty} \alpha_n(G) \geq \alpha(G)$$

Combining the two parts, if $A \subset \mathbf{R}$ is a continuity set of α i.e. $\alpha(\partial A) = \alpha(\bar{A} - A^o) = 0$, then

$$\lim_{n\to\infty} \alpha_n(A) = \alpha(A)$$

Exercise 2.6. If $\alpha_n \Rightarrow \alpha$ it is not always true that $\int x\, d\alpha_n \to \int x\, d\alpha$ because while x is a continuous function it is not bounded. Construct a simple counterexample. On the positive side, let $f(x)$ be a continuous function that is not necessarily bounded. Assume that there exists a positive continuous function $g(x)$ satisfying

$$\lim_{|x|\to\infty} \frac{|f(x)|}{g(x)} = 0$$

and

$$\sup_n \int g(x)\, d\alpha_n \leq C < \infty$$

Then show that

$$\lim_{n\to\infty} \int f(x)\, d\alpha_n = \int f(x)\, d\alpha$$

In particular if $\int |x|^k\, d\alpha_n$ remains bounded, then $\int |x|^j\, d\alpha_n \to \int |x|^j\, d\alpha$ for $1 \leq j \leq k-1$.