# Mean, Variance, Moments and Characteristic Functions

For a r.v *X*, its p.d.f $f_X(x)$ represents complete information about it, and for any Borel set *B* on the *x*-axis

$$P\big(X(\xi) \in B\big) = \int_B f_X(x)\,dx\,. \tag{6-1}$$

Note that $f_X(x)$ represents very detailed information, and quite often it is desirable to characterize the r.v in terms of its average behavior. In this context, we will introduce two parameters - mean and variance - that are universally used to represent the overall properties of the r.v and its p.d.f.

**Mean** or the **Expected Value** of a r.v $X$ is defined as

$$\eta_X = \overline{X} = E(X) = \int_{-\infty}^{+\infty} x \, f_X(x)dx. \tag{6-2}$$

If $X$ is a discrete-type r.v, then using (3-25) we get

$$\eta_X = \overline{X} = E(X) = \int x \sum_i p_i \delta(x - x_i)dx = \sum_i x_i p_i \underbrace{\int \delta(x - x_i)dx}_{1}$$

$$= \sum_i x_i p_i = \sum_i x_i \, P(X = x_i). \tag{6-3}$$

Mean represents the average (mean) value of the r.v in a very large number of trials. For example if $X \sim U(a, b)$, then using (3-31) ,

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \tag{6-4}$$

is the midpoint of the interval $(a,b)$.

On the other hand if $X$ is exponential with parameter $\lambda$ as in (3-32), then

$$E(X)=\int_0^{\infty}\frac{x}{\lambda}e^{-x/\lambda}dx=\lambda\int_0^{\infty}ye^{-y}dy=\lambda, \tag{6-5}$$

implying that the parameter $\lambda$ in (3-32) represents the mean value of the exponential r.v.

Similarly if X is Poisson with parameter $\lambda$ as in (3-45), using (6-3), we get

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} kP(X=k) = \sum_{k=0}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^k}{k!} \\ &= e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = \lambda e^{-\lambda}e^{\lambda} = \lambda. \end{aligned} \tag{6-6}$$

Thus the parameter $\lambda$ in (3-45) also represents the mean of the Poisson r.v.

In a similar manner, if $X$ is binomial as in (3-44), then its mean is given by

$$E(X)=\sum_{k=0}^{n}kP(X=k)=\sum_{k=0}^{n}k\binom{n}{k}p^k q^{n-k}=\sum_{k=1}^{n}k\frac{n!}{(n-k)!\,k!}p^k q^{n-k}$$
$$=\sum_{k=1}^{n}\frac{n!}{(n-k)!\,(k-1)!}p^k q^{n-k}=np\sum_{i=0}^{n-1}\frac{(n-1)!}{(n-i-1)!\,i!}p^i q^{n-i-1}=np(p+q)^{n-1}=np. \tag{6-7}$$

Thus $np$ represents the mean of the binomial r.v in (3-44).

For the normal r.v in (3-29),

$$E(X)=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{+\infty}xe^{-(x-\mu)^2/2\sigma^2}dx=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{+\infty}(y+\mu)e^{-y^2/2\sigma^2}dy$$
$$=\frac{1}{\sqrt{2\pi\sigma^2}}\underbrace{\int_{-\infty}^{+\infty}ye^{-y^2/2\sigma^2}dy}_{0}+\mu\cdot\underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{+\infty}e^{-y^2/2\sigma^2}dy}_{1}=\mu. \tag{6-8}$$

Thus the first parameter in $X \sim N(\mu, \sigma^2)$ is infact the mean of the Gaussian r.v *X*. Given $X \sim f_X(x)$, suppose $Y = g(X)$ defines a new r.v with p.d.f $f_Y(y)$. Then from the previous discussion, the new r.v *Y* has a mean $\mu_Y$ given by (see (6-2))

$$\mu_Y = E(Y) = \int_{-\infty}^{+\infty} y \; f_Y(y) dy . \tag{6-9}$$

From (6-9), it appears that to determine $E(Y)$, we need to determine $f_Y(y)$. However this is not the case if only $E(Y)$ is the quantity of interest. Recall that for any *y*, $\Delta y > 0$

$$P\big(y < Y \leq y + \Delta y\big) = \sum_i P\big(x_i < X \leq x_i + \Delta x_i\big), \tag{6-10}$$

where $x_i$ represent the multiple solutions of the equation $y = g(x_i)$. But(6-10) can be rewritten as

$$f_Y(y) \Delta y = \sum_i f_X(x_i) \Delta x_i , \tag{6-11}$$

where the $(x_i,\ x_i + \Delta x_i)$ terms form nonoverlapping intervals. Hence

$$y\ f_Y(y)\Delta y = \sum_i y\ f_X(x_i)\Delta x_i = \sum_i g(x_i)\, f_X(x_i)\Delta x_i, \tag{6-12}$$

and hence as *Δy* covers the entire y-axis, the corresponding *Δx*'s are nonoverlapping, and they cover the entire *x*-axis. Hence, in the limit as $\Delta y \to 0$, integrating both sides of (6-12), we get the useful formula

$$E(Y) = E\big(g(X)\big) = \int_{-\infty}^{+\infty} y\ f_Y(y)dy = \int_{-\infty}^{+\infty} g(x)\, f_X(x)dx. \tag{6-13}$$

In the discrete case, (6-13) reduces to

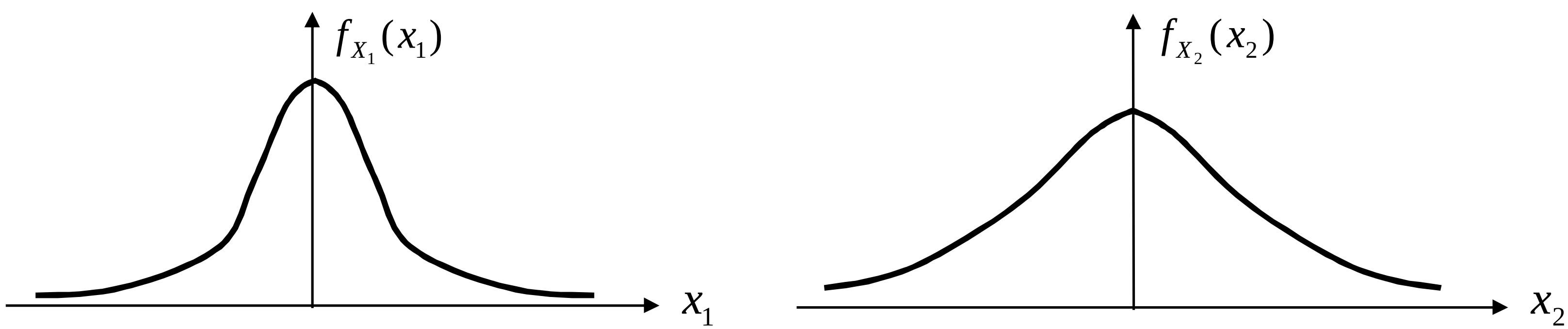
$$E(Y) = \sum_i g(x_i)P(X = x_i). \tag{6-14}$$

From (6-13)-(6-14), $f_Y(y)$ is not required to evaluate $E(Y)$ for $Y = g(X)$. We can use (6-14) to determine the mean of $Y = X^2,$ where *X* is a Poisson r.v. Using (3-45)

$$
\begin{aligned}
E\left(X^2\right) &= \sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k^2 \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \sum_{i=0}^{\infty} (i+1) \frac{\lambda^{i+1}}{i!} \\
&= \lambda e^{-\lambda} \left( \sum_{i=0}^{\infty} i \frac{\lambda^i}{i!} + \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \right) = \lambda e^{-\lambda} \left( \sum_{i=1}^{\infty} i \frac{\lambda^i}{i!} + e^{\lambda} \right) \\
&= \lambda e^{-\lambda} \left( \sum_{i=1}^{\infty} \frac{\lambda^i}{(i-1)!} + e^{\lambda} \right) = \lambda e^{-\lambda} \left( \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} + e^{\lambda} \right) \\
&= \lambda e^{-\lambda} \left( \lambda e^{\lambda} + e^{\lambda} \right) = \lambda^2 + \lambda. \qquad \text{(6-15)}
\end{aligned}
$$

In general, $E\left(X^k\right)$ is known as the *k*th moment of r.v *X*. Thus if $X \sim P(\lambda)$, its second moment is given by (6-15).

Mean alone will not be able to truly represent the p.d.f of any r.v. To illustrate this, consider the following scenario: Consider two Gaussian r.vs $X_1 \sim N(0,1)$ and $X_2 \sim N(0,10)$. Both of them have the same mean $\mu = 0$. However, as Fig. 6.1 shows, their p.d.fs are quite different. One is more concentrated around the mean, whereas the other one $(X_2)$ has a wider spread. Clearly, we need atleast an additional parameter to measure this spread around the mean!




**(a)** $\sigma^2 = 1$ (b) $\sigma^2 = 10$

Fig.6.1

For a r.v $X$ with mean $\mu$, $X - \mu$ represents the deviation of the r.v from its mean. Since this deviation can be either positive or negative, consider the quantity $(X - \mu)^2$, and its average value $E[(X - \mu)^2]$ represents the average mean square deviation of $X$ around its mean. Define

$$\sigma_X^2 \overset{\Delta}{=} E[(X - \mu)^2] > 0. \tag{6-16}$$

With $g(X) = (X - \mu)^2$ and using (6-13) we get

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f_X(x) dx > 0. \tag{6-17}$$

$\sigma_X^2$ is known as the variance of the r.v $X$, and its square root $\sigma_X = \sqrt{E(X - \mu)^2}$ is known as the standard deviation of $X$. Note that the standard deviation represents the root mean square spread of the r.v $X$ around its mean $\mu$.

Expanding (6-17) and using the linearity of the integrals, we get

$$Var(X) = \sigma_X^2 = \int_{-\infty}^{+\infty} \left( x^2 - 2x\mu + \mu^2 \right) f_X(x)dx$$

$$= \int_{-\infty}^{+\infty} x^2 f_X(x)dx - 2\mu \int_{-\infty}^{+\infty} x \ f_X(x)dx + \mu^2$$

$$= E\left(X^2\right) - \mu^2 = E\left(X^2\right) - \left[E(X)\right]^2 = \overline{X^2} - \overline{X}^2. \qquad \text{(6-18)}$$

Alternatively, we can use (6-18) to compute $\sigma_x^2$.

Thus , for example, returning back to the Poisson r.v in (3-45), using (6-6) and (6-15), we get

$$\sigma_x^2 = \overline{X^2} - \overline{X}^2 = \left(\lambda^2 + \lambda\right) - \lambda^2 = \lambda. \qquad \text{(6-19)}$$

Thus for a Poisson r.v, mean and variance are both equal to its parameter $\lambda$.

To determine the variance of the normal r.v $N(\mu, \sigma^2)$, we can use (6-16). Thus from (3-29)

$$Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx. \qquad (6\text{-}20)$$

To simplify (6-20), we can make use of the identity

$$\int_{-\infty}^{+\infty} f_X(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = 1$$

for a normal p.d.f. This gives

$$\int_{-\infty}^{+\infty} e^{-(x-\mu)^2/2\sigma^2} dx = \sqrt{2\pi}\,\sigma. \qquad (6\text{-}21)$$

Differentiating both sides of (6-21) with respect to $\sigma$, we get

$$\int_{-\infty}^{+\infty} \frac{(x-\mu)^2}{\sigma^3} e^{-(x-\mu)^2/2\sigma^2} dx = \sqrt{2\pi}$$

or

$$\int_{-\infty}^{+\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = \sigma^2, \qquad (6\text{-}22)$$

which represents the $Var(X)$ in (6-20). Thus for a normal r.v as in (3-29)

$$Var(X) = \sigma^2 \tag{6-23}$$

and the second parameter in $N(\mu, \sigma^2)$ infact represents the variance of the Gaussian r.v. As Fig. 6.1 shows the larger the $\sigma$, the larger the spread of the p.d.f around its mean. Thus as the variance of a r.v tends to zero, it will begin to concentrate more and more around the mean ultimately behaving like a constant.

**Moments**: As remarked earlier, in general

$$m_n = \overline{X^n} = E(X^n), \quad n \geq 1 \tag{6-24}$$

are known as the moments of the r.v *X*, and

$$\mu_n = E[(X-\mu)^n] \tag{6-25}$$

are known as the central moments of $X$. Clearly, the mean $\mu = m_1$, and the variance $\sigma^2 = \mu_2$. It is easy to relate $m_n$ and $\mu_n$. Infact

$$\mu_n = E[(X-\mu)^n] = E\left(\sum_{k=0}^{n}\binom{n}{k}X^k(-\mu)^{n-k}\right)$$

$$= \sum_{k=0}^{n}\binom{n}{k}E\left(X^k\right)(-\mu)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}\, m_k\,(-\mu)^{n-k}. \tag{6-26}$$

In general, the quantities

$$E[(X-a)^n] \tag{6-27}$$

are known as the generalized moments of $X$ about $a$, and

$$E[|X|^n] \tag{6-28}$$

are known as the absolute moments of $X$.

For example, if $X \sim N(0, \sigma^2)$, then it can be shown that

$$E(X^n) = \begin{cases} 0, & n \text{ odd}, \\ 1 \cdot 3 \cdots (n-1)\sigma^n, & n \text{ even}. \end{cases} \tag{6-29}$$

$$E(|X|^n) = \begin{cases} 1 \cdot 3 \cdots (n-1)\sigma^n, & n \text{ even}, \\ 2^k k! \sigma^{2k+1} \sqrt{2/\pi}, & n = (2k+1), \text{ odd}. \end{cases} \tag{6-30}$$

Direct use of (6-2), (6-13) or (6-14) is often a tedious procedure to compute the mean and variance, and in this context, the notion of the characteristic function can be quite helpful.

**Characteristic Function**

The characteristic function of a r.v $X$ is defined as

$$\Phi_X(\omega) \overset{\Delta}{=} E\left(e^{jX\omega}\right) = \int_{-\infty}^{+\infty} e^{jx\omega} f_X(x)dx. \tag{6-31}$$

Thus $\Phi_X(0) = 1,$ and $\left|\Phi_X(\omega)\right| \le 1$ for all $\omega.$

For discrete r.vs the characteristic function reduces to

$$\Phi_X(\omega) = \sum_k e^{jk\omega} P(X = k). \tag{6-32}$$

Thus for example, if $X \sim P(\lambda)$ as in (3-45), then its characteristic function is given by

$$\Phi_X(\omega) = \sum_{k=0}^{\infty} e^{jk\omega} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{j\omega})^k}{k!} = e^{-\lambda} e^{\lambda e^{j\omega}} = e^{\lambda(e^{j\omega}-1)}. \tag{6-33}$$

Similarly, if X is a binomial r.v as in (3-44), its characteristic function is given by

$$\Phi_X(\omega) = \sum_{k=0}^{n} e^{jk\omega} \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^{j\omega})^k q^{n-k} = (pe^{j\omega} + q)^n. \tag{6-34}$$

To illustrate the usefulness of the characteristic function of a r.v in computing its moments, first it is necessary to derive the relationship between them. Towards this, from (6-31)

$$\Phi_X(\omega) = E\left(e^{jX\omega}\right) = E\left[\sum_{k=0}^{\infty} \frac{(j\omega X)^k}{k!}\right] = \sum_{k=0}^{\infty} j^k \frac{E(X^k)}{k!}\omega^k$$

$$= 1 + jE(X)\omega + j^2 \frac{E(X^2)}{2!}\omega^2 + \cdots + j^k \frac{E(X^k)}{k!}\omega^k + \cdots. \qquad \text{(6-35)}$$

Taking the first derivative of (6-35) with respect to $\omega$, and letting it to be equal to zero, we get

$$\left.\frac{\partial \Phi_X(\omega)}{\partial \omega}\right|_{\omega=0} = jE(X) \quad \text{or} \quad E(X) = \frac{1}{j}\left.\frac{\partial \Phi_X(\omega)}{\partial \omega}\right|_{\omega=0}. \qquad \text{(6-36)}$$

Similarly, the second derivative of (6-35) gives

$$E(X^2) = \frac{1}{j^2}\left.\frac{\partial^2 \Phi_X(\omega)}{\partial \omega^2}\right|_{\omega=0}, \qquad \text{(6-37)}$$

and repeating this procedure *k* times, we obtain the *k*th moment of *X* to be

$$E(X^k) = \frac{1}{j^k} \left. \frac{\partial^k \Phi_X(\omega)}{\partial \omega^k} \right|_{\omega=0}, \quad k \geq 1. \tag{6-38}$$

We can use (6-36)-(6-38) to compute the mean, variance and other higher order moments of any r.v *X*. For example, if $X \sim P(\lambda)$, then from (6-33)

$$\frac{\partial \Phi_X(\omega)}{\partial \omega} = e^{-\lambda} e^{\lambda e^{j\omega}} \lambda j e^{j\omega}, \tag{6-39}$$

so that from (6-36)

$$E(X) = \lambda, \tag{6-40}$$

which agrees with (6-6). Differentiating (6-39) one more time, we get

$$\frac{\partial^2 \Phi_X(\omega)}{\partial \omega^2} = e^{-\lambda}\left(e^{\lambda e^{j\omega}}(\lambda j e^{j\omega})^2 + e^{\lambda e^{j\omega}} \lambda j^2 e^{j\omega}\right), \tag{6-41}$$

so that from (6-37)

$$E(X^2) = \lambda^2 + \lambda, \tag{6-42}$$

which again agrees with (6-15). Notice that compared to the tedious calculations in (6-6) and (6-15), the efforts involved in (6-39) and (6-41) are very minimal.

We can use the characteristic function of the binomial r.v $B(n, p)$ in (6-34) to obtain its variance. Direct differentiation of (6-34) gives

$$\frac{\partial \Phi_X(\omega)}{\partial \omega} = jnpe^{j\omega}(pe^{j\omega} + q)^{n-1} \tag{6-43}$$

so that from (6-36), $E(X) = np$ as in (6-7).

One more differentiation of (6-43) yields

$$\frac{\partial^2 \Phi_X(\omega)}{\partial \omega^2} = j^2 np\left(e^{j\omega}(pe^{j\omega}+q)^{n-1} + (n-1)pe^{j2\omega}(pe^{j\omega}+q)^{n-2}\right) \qquad (6\text{-}44)$$

and using (6-37), we obtain the second moment of the binomial r.v to be

$$E(X^2) = np\left(1+(n-1)p\right) = n^2p^2 + npq. \qquad (6\text{-}45)$$

Together with (6-7), (6-18) and (6-45), we obtain the variance of the binomial r.v to be

$$\sigma_X^2 = E(X^2) - \left[E(X)\right]^2 = n^2p^2 + npq - n^2p^2 = npq. \qquad (6\text{-}46)$$

To obtain the characteristic function of the Gaussian r.v, we can make use of (6-31). Thus if $X \sim N(\mu, \sigma^2)$, then

$$\Phi_X(\omega) = \int_{-\infty}^{+\infty} e^{j\omega x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx \quad \text{(Let } x - \mu = y\text{)}$$

$$= e^{j\mu\omega} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{j\omega y} e^{-y^2/2\sigma^2} dy = e^{j\mu\omega} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-y/2\sigma^2 (y - j2\sigma^2\omega)} dy$$

$$\text{(Let } y - j\sigma^2\omega = u \text{ so that } y = u + j\sigma^2\omega\text{)}$$

$$= e^{j\mu\omega} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-(u + j\sigma^2\omega)(u - j\sigma^2\omega)/2\sigma^2} du$$

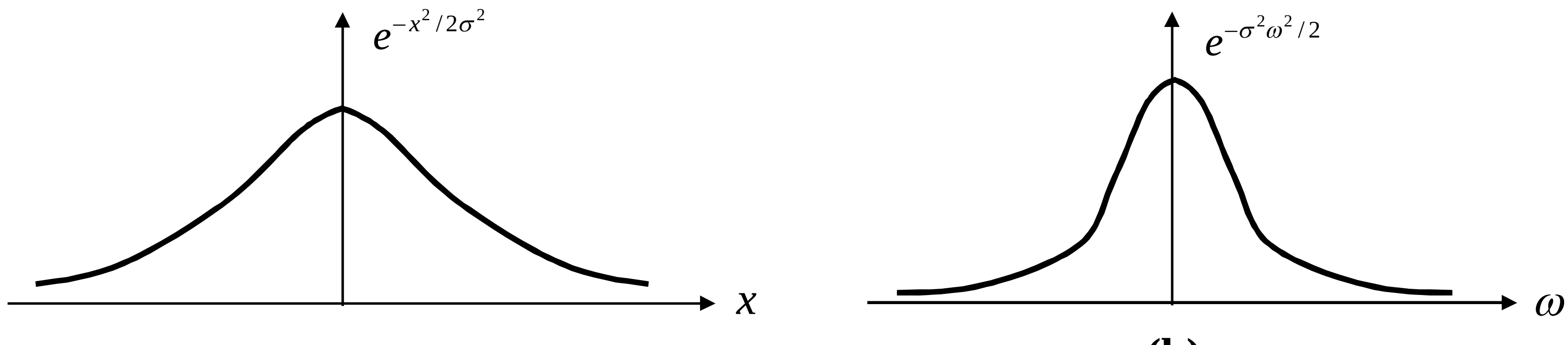
$$= e^{j\mu\omega} e^{-\sigma^2\omega^2/2} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{-u^2/2\sigma^2} du = e^{(j\mu\omega - \sigma^2\omega^2/2)}. \qquad (6\text{-}47)$$

Notice that the characteristic function of a Gaussian r.v itself has the "Gaussian" bell shape. Thus if $X \sim N(0, \sigma^2)$, then

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}, \qquad (6\text{-}48)$$

and

$$\Phi_X(\omega) = e^{-\sigma^2\omega^2/2}. \qquad (6\text{-}49)$$



(a) **(b)**

Fig. 6.2

From Fig. 6.2, the reverse roles of $\sigma^2$ in $f_X(x)$ and $\Phi_X(\omega)$ are noteworthy $(\sigma^2 \text{ vs } \frac{1}{\sigma^2})$.

In some cases, mean and variance may not exist. For example, consider the Cauchy r.v defined in (3-39). With

$f_X(x) = \frac{(\alpha/\pi)}{\alpha^2 + x^2}$,

$$E(X^2) = \frac{\alpha}{\pi}\int_{-\infty}^{+\infty}\frac{x^2}{\alpha^2 + x^2}dx = \frac{\alpha}{\pi}\int_{-\infty}^{+\infty}\left(1 - \frac{\alpha^2}{\alpha^2 + x^2}\right)dx = \infty, \qquad \text{(6-50)}$$

clearly diverges to infinity. Similarly

$$E(X) = \frac{\alpha}{\pi}\int_{-\infty}^{+\infty} \frac{x}{\alpha^2 + x^2}\,dx. \tag{6-51}$$

To compute (6-51), let us examine its one sided factor

$\int_0^{+\infty} \frac{x}{\alpha^2+x^2}\,dx.$ With $x = \alpha \tan\theta$

$$\int_0^{+\infty} \frac{x}{\alpha^2 + x^2}\,dx = \int_0^{\pi/2} \frac{\alpha \tan\theta}{\alpha^2 \sec^2\theta}\,\alpha \sec^2\theta\, d\theta = \int_0^{\pi/2} \frac{\sin\theta}{\cos\theta}\,d\theta$$

$$= -\int_0^{\pi/2} \frac{d(\cos\theta)}{\cos\theta} = -\log\cos\theta\Big|_0^{\pi/2} = -\log\cos\frac{\pi}{2} = \infty, \tag{6-52}$$

indicating that the double sided integral in (6-51) does not converge and is undefined. From (6-50)-(6-52), the mean and variance of a Cauchy r.v are undefined.

We conclude this section with a bound that estimates the dispersion of the r.v beyond a certain interval centered around its mean. Since $\sigma^2$ measures the dispersion of