3.2 Variances, Covariances and Correlation coefficients

3.2.1 Variances

3.2 Variances, Covariances and Correlation coefficients

3.2.1 Variances

Example 1. Let ξ and η denote the numbers of times two people, say A and B, hit the target respectively. Suppose

$$\xi : \begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix} \quad \eta : \begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

Compare which one is better at shooting.

3.2 Variances, Covariances and Correlation coefficients

3.2.1 Variances

Example 1. Let ξ and η denote the numbers of times two people, say A and B, hit the target respectively. Suppose

$$\xi : \begin{pmatrix} 7 & 8 & 9 \\ 0.1 & 0.8 & 0.1 \end{pmatrix} \quad \eta : \begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ 0.1 & 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}.$$

Compare which one is better at shooting.

$$E\xi = E\eta = 8.$$

We need consider the deviation extent to which it takes values, besides its mean for a random variable.

skip

We call $\xi - E\xi$ the deviation of ξ from its mean $E\xi$. It is still a random variable.

The mean of it is also $E[\xi - E\xi] = E\xi - E\xi = 0$.

Definition 1. If $E(\xi - E\xi)^2$ exists and is a finite constant, then we call it the variance of ξ, and write $Var\xi$ or $D\xi$, i.e.,

$$Var\xi = E(\xi - E\xi)^2.$$

Definition 1. If $E(\xi - E\xi)^2$ exists and is a finite constant, then we call it the variance of ξ, and write $Var\xi$ or $D\xi$, i.e.,

$$Var\xi = E(\xi - E\xi)^2.$$

But $Var\xi$ and ξ have different dimension. To unify, we sometimes use $\sqrt{Var\xi}$, called the standard deviation of ξ.

$$
\begin{aligned}
Var\xi &= \int_{-\infty}^{\infty}(x-E\xi)^2 dF_\xi(x) \\
&= \begin{cases} \sum_i (x_i - E\xi)^2 P(\xi = x_i) & \text{(discrete)}, \\ \int_{-\infty}^{\infty}(x-E\xi)^2 p_\xi(x)dx & \text{(continuous)}. \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
Var\xi &= \int_{-\infty}^{\infty} (x - E\xi)^2 dF_\xi(x) \\
&= \begin{cases} \sum_i (x_i - E\xi)^2 P(\xi = x_i) & \text{(discrete)}, \\ \int_{-\infty}^{\infty} (x - E\xi)^2 p_\xi(x)dx & \text{(continuous)}. \end{cases}
\end{aligned}
$$

Moreover,

$$E(\xi - E\xi)^2 = E[\xi^2 - 2\xi E\xi + (E\xi)^2] = E\xi^2 - (E\xi)^2,$$

i.e.,

$$Var\xi = E\xi^2 - (E\xi)^2.$$

Example 1 (continuity). Find $Var\xi$ and $Var\eta$ of ξ and η.

Solution.

Example 1 (continuity). Find $Var\xi$ and $Var\eta$ of ξ and η.

Solution.

$$E\xi^2 = \sum_i x_i^2 P(\xi = x_i) = 64.2,$$

so

$$Var\xi = E\xi^2 - (E\xi)^2 = 0.2.$$

Example 1 (continuity). Find $Var\xi$ and $Var\eta$ of ξ and η.

Solution.

$$E\xi^2 = \sum_i x_i^2 P(\xi = x_i) = 64.2,$$

so

$$Var\xi = E\xi^2 - (E\xi)^2 = 0.2.$$

Similarly,

$Var\eta = E\eta^2 - (E\eta)^2 = 65.2 - 64 = 1.2 > Var\xi$.

So η takes its values more dispersedly, which implies A shoots better.

Example 2. Find the variance of the Poisson distribution $P(\lambda)$.

Solution (1).

Example 2. Find the variance of the Poisson distribution $P(\lambda)$.

Solution (1).

$$E\xi^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda}$$

Example 2. Find the variance of the Poisson distribution $P(\lambda)$.

Solution (1).

$$
\begin{aligned}
E\xi^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda}
\end{aligned}
$$

Example 2. Find the variance of the Poisson distribution $P(\lambda)$.

Solution (1).

$$\begin{aligned} E\xi^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \lambda \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} + \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} = \lambda^2 + \lambda. \end{aligned}$$

Example 2. Find the variance of the Poisson distribution $P(\lambda)$.

Solution (1).

$$
\begin{aligned}
E\xi^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \sum_{k=1}^{\infty} (k-1) \frac{\lambda^k}{(k-1)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{j=0}^{\infty} j \frac{\lambda^j}{j!} e^{-\lambda} + \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{-\lambda} = \lambda^2 + \lambda.
\end{aligned}
$$

So, $Var\xi = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Solution (2).

Solution (2).

$$E\xi f(\xi-1)=\sum_{k=0}^{\infty}kf(k-1)\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\lambda\sum_{k=1}^{\infty}f(k-1)\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}=\lambda Ef(\xi).$$

Solution (2).

$$E\xi f(\xi-1)=\sum_{k=0}^{\infty}kf(k-1)\frac{\lambda^k}{k!}e^{-\lambda}$$

$$=\lambda\sum_{k=1}^{\infty}f(k-1)\frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}=\lambda Ef(\xi).$$

Letting $f(x)\equiv 1$ yields $E\xi=\lambda$. Letting $f(x)=x$ yields $E[\xi^2-\xi]=\lambda E\xi$. So

$$Var\xi=E\xi^2-(E\xi)^2=E\xi+\lambda E\xi-(E\xi)^2=\lambda.$$

Example 3. Suppose that $\xi \sim U[a, b]$, find $Var\xi$.

Solution.

Example 3. Suppose that $\xi \sim U[a,b]$, find $Var\xi$.

Solution. It is easy to see

$$E\xi^2 = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{1}{3}(a^2+ab+b^2).$$

Then

Example 3. Suppose that $\xi \sim U[a,b]$, find $Var\xi$.

Solution. It is easy to see

$$E\xi^2 = \int_a^b x^2 \frac{1}{b-a}dx = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{1}{3}(a^2+ab+b^2).$$

Then

$$Var\xi = \frac{1}{3}(a^2+ab+b^2) - [\frac{a+b}{2}]^2 = \frac{1}{12}(b-a)^2.$$

Example 4. Assume that $\xi \sim N(a, \sigma^2)$, find $Var\xi$.

Solution.

Example 4. Assume that $\xi \sim N(a, \sigma^2)$, find $Var\xi$.

Solution. Recall $E\xi = a$. We have

$$Var\xi = E(\xi - a)^2$$
$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$

Example 4. Assume that $\xi \sim N(a, \sigma^2)$, find $Var\xi$.

Solution. Recall $E\xi = a$. We have

$$\begin{aligned} Var\xi &= E(\xi - a)^2 \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} zd\left[-e^{-\frac{z^2}{2}}\right] \end{aligned}$$

Example 4. Assume that $\xi \sim N(a, \sigma^2)$, find $Var\xi$.

Solution. Recall $E\xi = a$. We have

$$
\begin{aligned}
& Var\xi = E(\xi - a)^2 \\
= \ & \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx \\
= \ & \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} zd\left[-e^{-\frac{z^2}{2}}\right] \\
= \ & \frac{\sigma^2}{\sqrt{2\pi}} (-ze^{-\frac{z^2}{2}}|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz) = \sigma^2.
\end{aligned}
$$

Chebyshev's inequality. For any $\varepsilon > 0$,

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

Proof.

Chebyshev's inequality. For any $\varepsilon > 0$,

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

Proof. Let

$$\eta = \begin{cases} 1, & \text{if } |\xi - E\xi| \geq \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\eta \leq \frac{|\xi - E\xi|^2}{\epsilon^2}.$$

Chebyshev's inequality. For any $\varepsilon > 0$,

$$P(|\xi - E\xi| \geq \varepsilon) \leq \frac{Var\xi}{\varepsilon^2}.$$

Proof. Let

$$\eta = \begin{cases} 1, & \text{if } |\xi - E\xi| \geq \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\eta \leq \frac{|\xi - E\xi|^2}{\epsilon^2}.$$

So

$$P(|\xi - E\xi| \geq \varepsilon) = E\eta \leq E\frac{|\xi - E\xi|^2}{\epsilon^2} = \frac{Var\xi}{\varepsilon^2}.$$

Properties of variances:

Properties of variances:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.

Properties of variances:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.

Proof. "$\Leftarrow$": Notice $E\xi = c$. So $P(\xi - E\xi = 0) = 1$. It follows that

$$Var(\xi) = E(\xi - E\xi)^2 = 0 * P(\xi - E\xi = 0) = 0.$$

Properties of variances:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.

Proof. "$\Leftarrow$": Notice $E\xi = c$. So $P(\xi - E\xi = 0) = 1$. It follows that

$$Var(\xi) = E(\xi - E\xi)^2 = 0 * P(\xi - E\xi = 0) = 0.$$

"$\Rightarrow$" For any $\epsilon > 0$,

$$P(|\xi - E\xi| \geq \epsilon) \leq \frac{Var(\xi)}{\epsilon^2} = 0.$$

Properties of variances:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.

Proof. "$\Leftarrow$": Notice $E\xi = c$. So $P(\xi - E\xi = 0) = 1$. It follows that

$$Var(\xi) = E(\xi - E\xi)^2 = 0 * P(\xi - E\xi = 0) = 0.$$

"$\Rightarrow$" For any $\epsilon > 0$,

$$P(|\xi - E\xi| \geq \epsilon) \leq \frac{Var(\xi)}{\epsilon^2} = 0.$$

Thus

$$P(\xi = E\xi) \quad = \quad 1 - P(|\xi - E\xi| > 0)$$

Properties of variances:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.

Proof. "$\Leftarrow$": Notice $E\xi = c$. So $P(\xi - E\xi = 0) = 1$. It follows that

$$Var(\xi) = E(\xi - E\xi)^2 = 0 * P(\xi - E\xi = 0) = 0.$$

"$\Rightarrow$" For any $\epsilon > 0$,

$$P(|\xi - E\xi| \geq \epsilon) \leq \frac{Var(\xi)}{\epsilon^2} = 0.$$

Thus

$$\begin{aligned} P(\xi = E\xi) &= 1 - P(|\xi - E\xi| > 0) \\ &= 1 - \lim_{n\to\infty} P\Big(|\xi - E\xi| \geq \frac{1}{n}\Big) = 1. \end{aligned}$$

2. $Var(c\xi + b) = c^2 Var\xi.$

2. $Var(c\xi + b) = c^2 Var\xi$.

Proof.

$$Var(c\xi + b) = E\left(c\xi + b - E(c\xi + b)\right)^2$$

2. $Var(c\xi + b) = c^2 Var\xi$.

Proof.

$$\begin{aligned} Var(c\xi + b) &= E\Big(c\xi + b - E(c\xi + b)\Big)^2 \\ &= E[c^2(\xi - E\xi)^2] \end{aligned}$$

2. $Var(c\xi + b) = c^2 Var\xi$.

Proof.

$$
\begin{aligned}
Var(c\xi + b) &= E\Big(c\xi + b - E(c\xi + b)\Big)^2 \\
&= E[c^2(\xi - E\xi)^2] \\
&= c^2 Var\xi.
\end{aligned}
$$

3. If $c \neq E\xi$, then $Var\xi < E(\xi - c)^2$.

3. If $c \neq E\xi$, then $Var\xi < E(\xi - c)^2$.

Proof.

$$
\begin{aligned}
& E(\xi - c)^2 \\
= \;& E\Big[\big((\xi - E\xi) + (E\xi - c)\big)^2\Big] \\
= \;& E\Big[(\xi - E\xi)^2 + 2(E\xi - c)(\xi - E\xi) + (E\xi - c)^2\Big]
\end{aligned}
$$

3. If $c \neq E\xi$, then $Var\xi < E(\xi - c)^2$.

Proof.

$$
\begin{aligned}
& E(\xi - c)^2 \\
= \;& E\Big[\big((\xi - E\xi) + (E\xi - c)\big)^2\Big] \\
= \;& E\Big[(\xi - E\xi)^2 + 2(E\xi - c)(\xi - E\xi) + (E\xi - c)^2\Big] \\
= \;& Var(\xi) + 2(E\xi - c)E[\xi - E\xi] + (E\xi - c)^2 \\
= \;& Var\xi + (E\xi - c)^2 \geq Var\xi.
\end{aligned}
$$

4

$$Var(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}Var\xi_i+2\sum_{1\le i<j\le n}E(\xi_i-E\xi_i)(\xi_j-E\xi_j).$$

If $\xi_1,\cdots,\xi_n$ are pairwise independent, then

$$Var(\sum_{i=1}^{n}\xi_i)=\sum_{i=1}^{n}Var\xi_i.$$

Proof.

$$Var(\sum_{i=1}^{n}\xi_i) = E(\sum_{i=1}^{n}\xi_i - E\sum_{i=1}^{n}\xi_i)^2$$

$$=$$

Proof.

$$Var(\sum_{i=1}^{n}\xi_i) = E(\sum_{i=1}^{n}\xi_i - E\sum_{i=1}^{n}\xi_i)^2$$

$$= E(\sum_{i=1}^{n}(\xi_i - E\xi_i))^2$$

$$=$$

Proof.

$$
\begin{aligned}
Var\Big(\sum_{i=1}^{n}\xi_i\Big) &= E\Big(\sum_{i=1}^{n}\xi_i - E\sum_{i=1}^{n}\xi_i\Big)^2 \\
&= E\Big(\sum_{i=1}^{n}(\xi_i - E\xi_i)\Big)^2 \\
&= E\Big[\sum_{i=1}^{n}(\xi_i - E\xi_i)^2 + 2\sum_{1\le i<j\le n}(\xi_i - E\xi_i)(\xi_j - E\xi_j)\Big] \\
&= \sum_{i=1}^{n} Var\xi_i + 2\sum_{1\le i<j\le n} E(\xi_i - E\xi_i)(\xi_j - E\xi_j).
\end{aligned}
$$

Summary:

1. $Var\xi = 0$ iff $P(\xi = c) = 1$.
2. $Var(c\xi + b) = c^2 Var\xi$.
3. If $c \neq E\xi$, then $Var\xi < E(\xi - c)^2$.
4. $$Var(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} Var\xi_i + 2 \sum_{1 \leq i < j \leq n} E(\xi_i - E\xi_i)(\xi_j - E\xi_j).$$

   If $\xi_1, \cdots, \xi_n$ are pairwise independent, then

   $$Var(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} Var\xi_i.$$

Example 5. Assume that $\xi \sim B(n, p)$, find $Var\xi$.

Solution.

Example 5. Assume that $\xi \sim B(n,p)$, find $Var\xi$.

Solution. Write $\xi = \xi_1 + \cdots + \xi_n$, where $\xi_1, \cdots, \xi_n$ i.i.d. $\sim B(1,p)$.

Example 5. Assume that $\xi \sim B(n,p)$, find $Var\xi$.

Solution. Write $\xi = \xi_1 + \cdots + \xi_n$, where $\xi_1, \cdots, \xi_n$ i.i.d. $\sim B(1,p)$. Then

$$Var\xi = \sum_{i=1}^{n} Var\xi_i = npq.$$

Example 6. Suppose that $\xi_1, \cdots, \xi_n$ are independent identically distributed random variables, and $E\xi_i = a, Var\xi_i = \sigma^2$. Let $\bar{\xi} = \sum_{i=1}^{n} \xi_i / n$, find $E\bar{\xi}$ and $Var\bar{\xi}$.

Example 6. Suppose that $\xi_1, \cdots, \xi_n$ are independent identically distributed random variables, and $E\xi_i = a, Var\xi_i = \sigma^2$. Let $\bar{\xi} = \sum_{i=1}^{n} \xi_i / n$, find $E\bar{\xi}$ and $Var\bar{\xi}$.

Solution.

$$E\bar{\xi} = \frac{1}{n} \sum_{i=1}^{n} E\xi_i = a;$$

Example 6. Suppose that $\xi_1, \cdots, \xi_n$ are independent identically distributed random variables, and $E\xi_i = a, Var\xi_i = \sigma^2$. Let $\bar{\xi} = \sum_{i=1}^n \xi_i/n$, find $E\bar{\xi}$ and $Var\bar{\xi}$.

Solution.

$$E\bar{\xi} = \frac{1}{n}\sum_{i=1}^n E\xi_i = a;$$

$$Var\bar{\xi} = \frac{1}{n^2}\sum_{i=1}^n Var\xi_i = \frac{\sigma^2}{n}.$$

Example 7. Suppose that the random variable ξ has finite expectation and positive variance. Let

$$\xi^* = \frac{\xi - E\xi}{\sqrt{Var\xi}}.$$

We call ξ^* the standardized random variable of ξ. Find $E\xi^*$ and $Var\xi^*$.

Example 7. Suppose that the random variable ξ has finite expectation and positive variance. Let

$$\xi^* = \frac{\xi - E\xi}{\sqrt{Var\xi}}.$$

We call ξ^* the standardized random variable of ξ. Find $E\xi^*$ and $Var\xi^*$.

Solution.

$$E\xi^* = \frac{E\xi - E\xi}{\sqrt{Var\xi}} = 0,$$

Example 7. Suppose that the random variable ξ has finite expectation and positive variance. Let

$$\xi^* = \frac{\xi - E\xi}{\sqrt{Var\xi}}.$$

We call ξ^* the standardized random variable of ξ. Find $E\xi^*$ and $Var\xi^*$.

Solution.

$$E\xi^* = \frac{E\xi - E\xi}{\sqrt{Var\xi}} = 0,$$

$$Var(\xi^*) = \frac{Var(\xi)}{Var(\xi)} = 1.$$

3.2.2 Covariances

For random vectors, say, $(\xi_1, \xi_2, \cdots, \xi_n)'$, besides expectation and variance of each coordinate, there is another numerical characteristic, called covariance, which expresses the connection between coordinate random variables.

Definition 2. Let $F_{ij}(x,y)$ be the joint distribution of ξ_i and ξ_j. If $E|(\xi_i - E\xi_i)(\xi_j - E\xi_j)| < \infty$, we call

$$E(\xi_i - E\xi_i)(\xi_j - E\xi_j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E\xi_i)(y - E\xi_j)dF_{ij}(x,y)$$

the covariance of ξ_i and ξ_j, written as $Cov(\xi_i, \xi_j)$.

Definition 2. Let $F_{ij}(x,y)$ be the joint distribution of ξ_i and ξ_j. If $E|(\xi_i - E\xi_i)(\xi_j - E\xi_j)| < \infty$, we call

$$E(\xi_i - E\xi_i)(\xi_j - E\xi_j)$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E\xi_i)(y - E\xi_j)dF_{ij}(x,y)$$

the covariance of ξ_i and ξ_j, written as $Cov(\xi_i, \xi_j)$.

$$Cov(\xi_i,\xi_j) = E(\xi_i - E\xi_i)(\xi_j - E\xi_j) = E\xi_i\xi_j - E\xi_i E\xi_j.$$

$$Cov(\xi_i, \xi_i) = Var\xi_i.$$

$$Var\left(\sum_{i=1}^{n} \xi_i\right) = \sum_{i=1}^{n} Var\xi_i + 2 \sum_{1 \le i < j \le n} Cov(\xi_i, \xi_j).$$

Properties of Covariances:

Properties of Covariances:

1. $Cov(\xi, \eta) = Cov(\eta, \xi) = E\xi\eta - E\xi E\eta;$

Properties of Covariances:

1. $Cov(\xi,\eta) = Cov(\eta,\xi) = E\xi\eta - E\xi E\eta$;
2. Assume that a, b are constants, then

$$Cov(a\xi, b\eta) = abCov(\xi,\eta);$$

Properties of Covariances:

1. $Cov(\xi,\eta)=Cov(\eta,\xi)=E\xi\eta-E\xi E\eta$;
2. Assume that a,b are constants, then
$$Cov(a\xi,b\eta)=abCov(\xi,\eta);$$
3. $Cov(\sum_{i=1}^n\xi_i,\eta)=\sum_{i=1}^n Cov(\xi_i,\eta)$.

Given an n-dimensional random vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)'$, one can write its covariance matrix as

$$\begin{aligned} B &= Var(\boldsymbol{\xi}) := E(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})' \\ &= \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}, \end{aligned}$$

where $b_{ij} = Cov(\xi_i, \xi_j)$.

From Property 1 it follows that B is a symmetric matrix and for any real numbers $t_j, j = 1, 2, \cdots, n$,

$$\begin{aligned}\sum_{j,k} b_{jk} t_j t_k &= \sum_{j,k} t_j t_k E(\xi_j - E\xi_j)(\xi_k - E\xi_k) \\ &= E\Big(\sum_{j=1}^{n} t_j(\xi_j - E\xi_j)\Big)^2 \geq 0,\end{aligned}$$

that is, the covariance matrix B is non-negative definite.

4. Let

$$\begin{aligned} \boldsymbol{\xi} &= (\xi_1, \xi_2, \cdots, \xi_n)', \\ C &= \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ C_{m1} & C_{m2} & \cdots & C_{mn} \end{pmatrix}, \end{aligned}$$

then $C\boldsymbol{\xi}$ has covariance matrix CBC', where B is covariance matrix of $\boldsymbol{\xi}$.

Indeed, it is easy to see

$$E\left[C(\boldsymbol{\xi}-E\boldsymbol{\xi})(C(\boldsymbol{\xi}-E\boldsymbol{\xi}))'\right]$$

Indeed, it is easy to see

$$\begin{aligned} & E\left[C(\boldsymbol{\xi} - E\boldsymbol{\xi})(C(\boldsymbol{\xi} - E\boldsymbol{\xi}))'\right] \\ = \; & E\left[C(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})'C'\right] \end{aligned}$$

Indeed, it is easy to see

$$\begin{aligned} & E\left[C(\xi - E\xi)(C(\xi - E\xi))'\right] \\ = \ & E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \\ = \ & E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \end{aligned}$$

Indeed, it is easy to see

$$
\begin{aligned}
& E\left[C(\xi - E\xi)(C(\xi - E\xi))'\right] \\
= \; & E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \\
= \; & E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \\
= \; & CE\left[(\xi - E\xi)(\xi - E\xi)'C'\right]
\end{aligned}
$$

Indeed, it is easy to see

$$
\begin{aligned}
& E\left[C(\xi-E\xi)(C(\xi-E\xi))'\right] \\
= & E\left[C(\xi-E\xi)(\xi-E\xi)'C'\right] \\
= & E\left[C(\xi-E\xi)(\xi-E\xi)'C'\right] \\
= & CE\left[(\xi-E\xi)(\xi-E\xi)'C'\right] \\
= & CE\left[(\xi-E\xi)(\xi-E\xi)'\right]C'
\end{aligned}
$$

Indeed, it is easy to see

$$\begin{aligned}
& E\left[C(\xi - E\xi)(C(\xi - E\xi))'\right] \\
= \;& E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \\
= \;& E\left[C(\xi - E\xi)(\xi - E\xi)'C'\right] \\
= \;& CE\left[(\xi - E\xi)(\xi - E\xi)'C'\right] \\
= \;& CE\left[(\xi - E\xi)(\xi - E\xi)'\right]C' \\
= \;& CBC'.
\end{aligned}$$

Indeed, it is easy to see

$$\begin{aligned} & E\left[C(\boldsymbol{\xi}-E\boldsymbol{\xi})(C(\boldsymbol{\xi}-E\boldsymbol{\xi}))'\right] \\ = & E\left[C(\boldsymbol{\xi}-E\boldsymbol{\xi})(\boldsymbol{\xi}-E\boldsymbol{\xi})'C'\right] \\ = & E\left[C(\boldsymbol{\xi}-E\boldsymbol{\xi})(\boldsymbol{\xi}-E\boldsymbol{\xi})'C'\right] \\ = & CE\left[(\boldsymbol{\xi}-E\boldsymbol{\xi})(\boldsymbol{\xi}-E\boldsymbol{\xi})'C'\right] \\ = & CE\left[(\boldsymbol{\xi}-E\boldsymbol{\xi})(\boldsymbol{\xi}-E\boldsymbol{\xi})'\right]C' \\ = & CBC'. \end{aligned}$$

Hence the (i,j)-th entry in CBC' is just the covariance of the i-th entry and the j-th entry in $C\boldsymbol{\xi}$.

5. If $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$E\boldsymbol{\xi} = \boldsymbol{\mu}, \quad Var(\boldsymbol{\xi}) = \boldsymbol{\Sigma}.$$

5. If $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$E\boldsymbol{\xi} = \boldsymbol{\mu}, \quad Var(\boldsymbol{\xi}) = \boldsymbol{\Sigma}.$$

Proof. First, consider the special case that $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_{n\times n}$. In this case, the pdf of $\boldsymbol{\xi}$ is

$$p(\boldsymbol{y}) = \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y_i^2}{2}\},$$

5. If $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then

$$E\boldsymbol{\xi} = \boldsymbol{\mu}, \quad Var(\boldsymbol{\xi}) = \boldsymbol{\Sigma}.$$

Proof. First, consider the special case that $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_{n\times n}$. In this case, the pdf of $\boldsymbol{\xi}$ is

$$p(\boldsymbol{y}) = \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{y_i^2}{2}\},$$

which means that $\xi_1, \cdots, \xi_n$ are i.i.d. $N(0,1)$ random variables. Hence

$$E\boldsymbol{\xi} = \mathbf{0}, \quad Var(\boldsymbol{\xi}) = \boldsymbol{I}_{n\times n}.$$

Now consider the general case. Since $\boldsymbol{\Sigma}$ is positive definite, there is a non-singular matrix $\boldsymbol{L} = \boldsymbol{\Sigma}^{1/2}$ such that $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}'$. Let $\boldsymbol{\eta} = \boldsymbol{L}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu})$ and notice $|\boldsymbol{L}| = |\boldsymbol{\Sigma}|^{1/2}$. Then $\boldsymbol{\xi} = \boldsymbol{L}\boldsymbol{\eta} + \boldsymbol{\mu}$, and the density of $\boldsymbol{\eta}$ is

$$p_{\eta}(\boldsymbol{y}) = p(\boldsymbol{x})|\boldsymbol{L}| \qquad (\boldsymbol{x} = \boldsymbol{L}\boldsymbol{y} + \boldsymbol{\mu})$$

Now consider the general case. Since $\boldsymbol{\Sigma}$ is positive definite, there is a non-singular matrix $\boldsymbol{L} = \boldsymbol{\Sigma}^{1/2}$ such that $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}'$. Let $\boldsymbol{\eta} = \boldsymbol{L}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu})$ and notice $|\boldsymbol{L}| = |\boldsymbol{\Sigma}|^{1/2}$. Then $\boldsymbol{\xi} = \boldsymbol{L}\boldsymbol{\eta} + \boldsymbol{\mu}$, and the density of $\boldsymbol{\eta}$ is

$$\begin{aligned} p_{\eta}(\boldsymbol{y}) &= p(\boldsymbol{x})|\boldsymbol{L}| \qquad\qquad (\boldsymbol{x} = \boldsymbol{L}\boldsymbol{y} + \boldsymbol{\mu}) \\ &= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\}|\boldsymbol{L}| \\ &= \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\}, \end{aligned}$$

Now consider the general case. Since $\boldsymbol{\Sigma}$ is positive definite, there is a non-singular matrix $\boldsymbol{L} = \boldsymbol{\Sigma}^{1/2}$ such that $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}'$. Let $\boldsymbol{\eta} = \boldsymbol{L}^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu})$ and notice $|\boldsymbol{L}| = |\boldsymbol{\Sigma}|^{1/2}$. Then $\boldsymbol{\xi} = \boldsymbol{L}\boldsymbol{\eta} + \boldsymbol{\mu}$, and the density of $\boldsymbol{\eta}$ is

$$\begin{aligned} p_{\eta}(\boldsymbol{y}) &= p(\boldsymbol{x})|\boldsymbol{L}| \qquad\qquad (\boldsymbol{x} = \boldsymbol{L}\boldsymbol{y} + \boldsymbol{\mu}) \\ &= \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}|\boldsymbol{L}| \\ &= \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\}, \end{aligned}$$

which means that $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{I}_{n\times n})$. So, $E\boldsymbol{\eta} = \mathbf{0}$ and $Var(\boldsymbol{\eta}) = \boldsymbol{I}$.

Hence

$$E\boldsymbol{\xi} = \boldsymbol{L}E\boldsymbol{\eta} + \boldsymbol{\mu} = \boldsymbol{\mu}.$$

and

Hence

$$E\boldsymbol{\xi} = \boldsymbol{L}E\boldsymbol{\eta} + \boldsymbol{\mu} = \boldsymbol{\mu}.$$

and

$$\begin{aligned} Var(\boldsymbol{\xi}) &= E[(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})'] \\ &= E\left[\boldsymbol{L}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \end{aligned}$$

Hence

$$E\boldsymbol{\xi} = \boldsymbol{L}E\boldsymbol{\eta} + \boldsymbol{\mu} = \boldsymbol{\mu}.$$

and

$$\begin{aligned} Var(\boldsymbol{\xi}) &= E[(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})'] \\ &= E\left[\boldsymbol{L}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \\ &= \boldsymbol{L}E\left[\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \end{aligned}$$

Hence

$$E\boldsymbol{\xi} = \boldsymbol{L}E\boldsymbol{\eta} + \boldsymbol{\mu} = \boldsymbol{\mu}.$$

and

$$\begin{aligned} Var(\boldsymbol{\xi}) &= E[(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})'] \\ &= E\left[\boldsymbol{L}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \\ &= \boldsymbol{L}E\left[\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \\ &= \boldsymbol{L}E\left[\boldsymbol{\eta}\boldsymbol{\eta}'\right]\boldsymbol{L}' \end{aligned}$$

Hence

$$E\boldsymbol{\xi} = \boldsymbol{L}E\boldsymbol{\eta} + \boldsymbol{\mu} = \boldsymbol{\mu}.$$

and

$$\begin{aligned} Var(\boldsymbol{\xi}) &= E[(\boldsymbol{\xi} - E\boldsymbol{\xi})(\boldsymbol{\xi} - E\boldsymbol{\xi})'] \\ &= E\left[\boldsymbol{L}\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \\ &= \boldsymbol{L}E\left[\boldsymbol{\eta}\boldsymbol{\eta}'\boldsymbol{L}'\right] \\ &= \boldsymbol{L}E\left[\boldsymbol{\eta}\boldsymbol{\eta}'\right]\boldsymbol{L}' \\ &= \boldsymbol{L}\boldsymbol{I}\boldsymbol{L}' = \boldsymbol{\Sigma}. \end{aligned}$$

3.2.3 Correlation coefficients

3.2.3 Correlation coefficients

Definition 3. We call

$$r_{\xi\eta} = Cov(\xi^*, \eta^*) = \frac{E(\xi - E\xi)(\eta - E\eta)}{\sqrt{Var\xi Var\eta}}$$

the correlation coefficient of ξ and η, where

$$\xi^* = \frac{\xi - E\xi}{\sqrt{Var\xi}}$$

$$\eta^* = \frac{\eta - E\eta}{\sqrt{Var\eta}}.$$

Cauchy-Schwarz's inequality. For any pair of random variables ξ and η with $P(\xi = 0) \neq 1$ and $P(\eta = 0) \neq 1$, it holds that

$$|E\xi\eta|^2 \leq E\xi^2 E\eta^2.$$

and the equality is valid if and only if there is a constant t_0 such that

$$P(\eta = t_0\xi) = 1.$$

Proof. Define

$$u(t) = E(t\xi - \eta)^2 = t^2 E\xi^2 - 2tE\xi\eta + E\eta^2.$$

Proof. Define

$$u(t) = E(t\xi - \eta)^2 = t^2 E\xi^2 - 2tE\xi\eta + E\eta^2.$$

This can be viewed as a non-negative quadratic in t, so its discriminant is

$$(E\xi\eta)^2 - E\xi^2 E\eta^2 \le 0.$$

Proof. Define

$$u(t) = E(t\xi - \eta)^2 = t^2 E\xi^2 - 2tE\xi\eta + E\eta^2.$$

This can be viewed as a non-negative quadratic in t, so its discriminant is

$$(E\xi\eta)^2 - E\xi^2 E\eta^2 \leq 0.$$

The equality is valid iff $u(t)$ has a multi-root $t_0 = E\xi\eta / E\xi^2$. In other words,

$$u(t_0) = E(t_0\xi - \eta)^2 = 0,$$

which implies $P(t_0\xi - \eta = 0) = 1$.

1. Let $r_{\xi\eta}$ be the correlation coefficient, then

$$|r_{\xi\eta}| \leq 1.$$

Also, $r_{\xi\eta} = 1$ if and only if

$$P\left(\frac{\xi - E\xi}{\sqrt{Var\xi}} = \frac{\eta - E\eta}{\sqrt{Var\eta}}\right) = 1;$$

$r_{\xi\eta} = -1$ if and only if

$$P\left(\frac{\xi - E\xi}{\sqrt{Var\xi}} = -\frac{\eta - E\eta}{\sqrt{Var\eta}}\right) = 1.$$

Proof. It follows that

$$|r_{\xi\eta}| = |E\xi^*\eta^*|$$

Proof. It follows that

$$\begin{aligned} |r_{\xi\eta}| &= |E\xi^*\eta^*| \leq \sqrt{E\xi^{*2}E\eta^{*2}} \\ &= \sqrt{Var\xi^* Var\eta^*} = 1. \end{aligned}$$

Proof. It follows that

$$\begin{aligned} |r_{\xi\eta}| &= |E\xi^*\eta^*| \leq \sqrt{E\xi^{*2}E\eta^{*2}} \\ &= \sqrt{Var\xi^* Var\eta^*} = 1. \end{aligned}$$

Next let us turn to the second conclusion. By the definition of correlation coefficient,

$r_{\xi\eta} = r_{\xi^*\eta^*} = E[\xi^*\eta^*]$.

Proof. It follows that

$$\begin{aligned} |r_{\xi\eta}| &= |E\xi^*\eta^*| \leq \sqrt{E\xi^{*2}E\eta^{*2}} \\ &= \sqrt{Var\xi^* Var\eta^*} = 1. \end{aligned}$$

Next let us turn to the second conclusion. By the definition of correlation coefficient, $r_{\xi\eta} = r_{\xi^*\eta^*} = E[\xi^*\eta^*]$. By the Cauchy-Schwarz's inequality, $|r_{\xi\eta}| = 1$ if and only if there exists t_0 such that $P(\eta^* = t_0\xi^*) = 1$.

It follows that

$$r_{\xi\eta} = E[\xi^*\eta^*] = E[t_0(\xi^*)^2] = t_0 Var\xi^* = t_0.$$

It follows that

$$r_{\xi\eta} = E[\xi^*\eta^*] = E[t_0(\xi^*)^2] = t_0 Var\xi^* = t_0.$$

Therefore $r_{\xi\eta} = 1$ iff $P(\xi^* = \eta^*) = 1$, while $r_{\xi\eta} = -1$ iff $P(\xi^* = -\eta^*) = 1$.

When $r_{\xi\eta} = 0$, we say ξ and η are uncorrelated.

When $r_{\xi\eta} = 0$, we say ξ and η are uncorrelated.

2. The following statements are equivalent
   1. $Cov(\xi, \eta) = 0$;
   2. ξ and η are uncorrelated;
   3. $E\xi\eta = E\xi E\eta$;
   4. $Var(\xi + \eta) = Var\xi + Var\eta$.

   Proof?

When $r_{\xi\eta} = 0$, we say ξ and η are uncorrelated.

2. The following statements are equivalent
   1. $Cov(\xi, \eta) = 0$;
   2. ξ and η are uncorrelated;
   3. $E\xi\eta = E\xi E\eta$;
   4. $Var(\xi + \eta) = Var\xi + Var\eta$.

   Proof?

3. If ξ and η are independent and their variances are finite, then ξ and η are uncorrelated.
   proof. trivial

$$\text{independence} \quad \not\Leftarrow \quad \text{uncorrelation}$$

independence $\not\Leftarrow$ uncorrelation

Example 8. $\theta \sim [0, 2\pi]$. Let $\xi = \cos\theta$, $\eta = \sin\theta$. Since $\xi^2 + \eta^2 = 1$, ξ, η are not indept.. However ξ, η are uncorrelated. Indeed,

$$E\xi = E\cos\theta = \int_0^{2\pi} \frac{1}{2\pi} \cos\varphi d\varphi = 0,$$

$$\text{independence} \not\Leftarrow \text{uncorrelation}$$

Example 8. $\theta \sim [0, 2\pi]$. Let $\xi = \cos\theta$, $\eta = \sin\theta$. Since $\xi^2 + \eta^2 = 1$, ξ, η are not indept.. However ξ, η are uncorrelated. Indeed,

$$\begin{aligned} E\xi &= E\cos\theta = \int_0^{2\pi} \frac{1}{2\pi} \cos\varphi d\varphi = 0, \\ E\eta &= E\sin\theta = \int_0^{2\pi} \frac{1}{2\pi} \sin\varphi d\varphi = 0, \end{aligned}$$

independence $\not\Leftarrow$ uncorrelation

Example 8. $\theta \sim [0, 2\pi]$. Let $\xi = \cos\theta$, $\eta = \sin\theta$. Since $\xi^2 + \eta^2 = 1$, ξ, η are not indept.. However ξ, η are uncorrelated. Indeed,

$$
\begin{aligned}
E\xi &= E\cos\theta = \int_0^{2\pi} \frac{1}{2\pi} \cos\varphi d\varphi = 0, \\
E\eta &= E\sin\theta = \int_0^{2\pi} \frac{1}{2\pi} \sin\varphi d\varphi = 0, \\
E\xi\eta &= E\sin\theta\cos\theta = \int_0^{2\pi} \frac{1}{2\pi} \sin\varphi\cos\varphi d\varphi = 0
\end{aligned}
$$

Thus $Cov(\xi, \eta) = E\xi\eta - E\xi E\eta = 0$.

In the case of normal distribution,

$$\text{independence} \iff \text{uncorrelation}$$

In the case of normal distribution,

independence $\Leftrightarrow$ uncorrelation

Example 9. Assume that $\xi, \eta \sim N(a, b, \sigma_1^2, \sigma_2^2, r)$, find $Cov(\xi, \eta)$ and $r_{\xi,\eta}$.

Solution (1).

$$
\begin{aligned}
& Cov(\xi, \eta) \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-a)(y-b)p(x,y)dxdy
\end{aligned}
$$

Solution (1).

$$
\begin{aligned}
& Cov(\xi, \eta) \\
= & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-a)(y-b)p(x,y)dxdy \\
= & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-a)(y-b) \\
& \cdot \exp(-\frac{1}{2(1-r^2)}(\frac{x-a}{\sigma_1}-r\frac{y-b}{\sigma_2})^2-\frac{(y-b)^2}{2\sigma_2^2})dxdy.
\end{aligned}
$$

Let

$$
z=\frac{x-a}{\sigma_1}-r\frac{y-b}{\sigma_2}, \quad t=\frac{y-b}{\sigma_2},
$$

then

$$
\frac{x-a}{\sigma_1}=z+rt, \qquad J=\frac{\partial(x,y)}{\partial(z,t)}=\sigma_1\sigma_2.
$$

Thus we have

$$
\begin{aligned}
& Cov(\xi, \eta) \\
= \quad & \frac{\sigma_1 \sigma_2}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (zt + rt^2) e^{-\frac{z^2}{2(1-r^2)}} e^{-\frac{t^2}{2}} dz dt
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
& Cov(\xi,\eta) \\
= \ & \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(zt+rt^2)e^{-\frac{z^2}{2(1-r^2)}}e^{-\frac{t^2}{2}}dzdt \\
= \ & \sigma_1\sigma_2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2(1-r^2)}}dz \\
& +\frac{r\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2(1-r^2)}}dz
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
& Cov(\xi,\eta) \\
= \;& \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(zt+rt^2)e^{-\frac{z^2}{2(1-r^2)}}e^{-\frac{t^2}{2}}dzdt \\
= \;& \sigma_1\sigma_2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2(1-r^2)}}dz \\
& +\frac{r\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2(1-r^2)}}dz \\
= \;& r\sigma_1\sigma_2.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
& Cov(\xi,\eta) \\
= \; & \frac{\sigma_1\sigma_2}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(zt+rt^2)e^{-\frac{z^2}{2(1-r^2)}}e^{-\frac{t^2}{2}}dzdt \\
= \; & \sigma_1\sigma_2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2(1-r^2)}}dz \\
& +\frac{r\sigma_1\sigma_2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t^2e^{-\frac{t^2}{2}}dt\frac{1}{\sqrt{2\pi}\sqrt{1-r^2}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2(1-r^2)}}dz \\
= \; & r\sigma_1\sigma_2.
\end{aligned}
$$

Therefore

$$
r_{\xi\eta}=\frac{Cov(\xi,\eta)}{\sqrt{Var\xi Var\eta}}=r.
$$

Solution (2). Notice

$$Cov(\xi,\eta) = E\left[(\xi-a)(\eta-b)\right] = E\left[E\left[(\xi-a)(\eta-b)|\xi\right]\right].$$

Solution (2). Notice

$$Cov(\xi,\eta) = E\left[(\xi-a)(\eta-b)\right] = E\left[E\left[(\xi-a)(\eta-b)|\xi\right]\right].$$

From

$$\eta|_{\xi=x} \sim N\Big(b + r\frac{\sigma_2}{\sigma_1}(x-a), (1-r^2)\sigma_2^2\Big),$$

it follows that

$$\begin{aligned} E\left[(\xi-a)(\eta-b)|\xi=x\right] =&(x-a)E\left[(\eta-b)|\xi=x\right] \\ =&r\frac{\sigma_2}{\sigma_1}(x-a)^2. \end{aligned}$$

Hence

$$E\left[(\xi-a)(\eta-b)|\xi\right] = r\frac{\sigma_2}{\sigma_1}(\xi-a)^2.$$

So

$$Cov(\xi, \eta) = E\left[r\frac{\sigma_2}{\sigma_1}(\xi - a)^2\right]$$

So

$$\begin{aligned} Cov(\xi,\eta) &= E\left[r\frac{\sigma_2}{\sigma_1}(\xi-a)^2\right] \\ &= r\frac{\sigma_2}{\sigma_1}E(\xi-a)^2 \end{aligned}$$

So

$$\begin{aligned} Cov(\xi,\eta) &= E\left[r\frac{\sigma_2}{\sigma_1}(\xi-a)^2\right] \\ &= r\frac{\sigma_2}{\sigma_1}E(\xi-a)^2 = r\sigma_1\sigma_2. \end{aligned}$$

So

$$\begin{aligned} Cov(\xi,\eta) &= E\left[r\frac{\sigma_2}{\sigma_1}(\xi-a)^2\right] \\ &= r\frac{\sigma_2}{\sigma_1}E(\xi-a)^2 = r\sigma_1\sigma_2. \end{aligned}$$

Therefore $r_{\xi\eta} = r$.

Solution (3). $(\xi,\eta)' \sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = (a,b)', \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

By Property 5,

$$Var\{(\xi,\eta)'\} = \boldsymbol{\Sigma},$$

i.e., $Var\xi = \sigma_1^2$, $Var\eta = \sigma_2^2$, $Cov(\xi,\eta) = r\sigma_1\sigma_2$. It follows that

$$r_{\xi\eta} = \frac{Cov(\xi,\eta)}{\sqrt{Var\xi Var\eta}} = r.$$

If (ξ, η) follows a normal distribution, then

$$\begin{aligned} &\xi, \eta \text{ are uncorrelated} \\ &\Leftrightarrow r_{\xi\eta} = 0 \\ &\Leftrightarrow r = 0 \\ &\Leftrightarrow \xi, \eta \text{ are indept.} \end{aligned}$$

4. For a bivariate normal distribution the uncorrelated property is equivalent to the independence.

In general, the random variables $\xi_1, \cdots, \xi_n$ with joint normal distribution are mutually independent iff they are pairwise uncorrelated.

Proof.

In general, the random variables $\xi_1, \cdots, \xi_n$ with joint normal distribution are mutually independent iff they are pairwise uncorrelated.

Proof. Assume $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

In general, the random variables $\xi_1, \cdots, \xi_n$ with joint normal distribution are mutually independent iff they are pairwise uncorrelated.

Proof. Assume $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then $Var(\boldsymbol{\xi}) = \boldsymbol{\Sigma}$, i.e., $Cov(\xi_i, \xi_j) = \sigma_{ij}$. So

$\xi_1, \cdots, \xi_n$ are indept.

$$\begin{aligned} & \xi_1, \cdots, \xi_n \text{ are indept.} \\ \Longrightarrow \quad & \xi_1, \cdots, \xi_n \text{ pairwise uncorrelated} \end{aligned}$$

$$\begin{aligned} & \xi_1, \cdots, \xi_n \text{ are indept.} \\ \Longrightarrow \quad & \xi_1, \cdots, \xi_n \text{ pairwise uncorrelated} \\ \Longrightarrow \quad & \sigma_{i,j} = Cov(\xi_i, \xi_j) = 0, \ i \neq j \end{aligned}$$

$$
\begin{aligned}
& \xi_1, \cdots, \xi_n \text{ are indept.} \\
\Longrightarrow\quad & \xi_1, \cdots, \xi_n \text{ pairwise uncorrelated} \\
\Longrightarrow\quad & \sigma_{i,j} = Cov(\xi_i, \xi_j) = 0,\ i \neq j \\
\Longrightarrow\quad & \mathbf{\Sigma} = diag(\sigma_1^2, \cdots, \sigma_n^2) \\
\Longrightarrow\quad & \mathbf{\Sigma}^{-1} = diag\left(\frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_n^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \xi_1, \cdots, \xi_n \text{ are indept.} \\
\Longrightarrow \quad & \xi_1, \cdots, \xi_n \text{ pairwise uncorrelated} \\
\Longrightarrow \quad & \sigma_{i,j} = Cov(\xi_i, \xi_j) = 0, \ i \neq j \\
\Longrightarrow \quad & \boldsymbol{\Sigma} = diag(\sigma_1^2, \cdots, \sigma_n^2) \\
\Longrightarrow \quad & \boldsymbol{\Sigma}^{-1} = diag\Big(\frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_n^2}\Big) \\
\Longrightarrow \quad & p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} \exp\Big\{ -\frac{1}{2} \sum_{i=1}^n \frac{(x_i - \mu_i)^2}{\sigma_i^2} \Big\} \\
& \qquad = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\{-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}\}
\end{aligned}
$$

$$
\begin{aligned}
& \xi_1, \cdots, \xi_n \text{ are indept.} \\
\Longrightarrow\quad & \xi_1, \cdots, \xi_n \text{ pairwise uncorrelated} \\
\Longrightarrow\quad & \sigma_{i,j} = Cov(\xi_i, \xi_j) = 0,\ i \neq j \\
\Longrightarrow\quad & \boldsymbol{\Sigma} = diag(\sigma_1^2, \cdots, \sigma_n^2) \\
\Longrightarrow\quad & \boldsymbol{\Sigma}^{-1} = diag\Big(\frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_n^2}\Big) \\
\Longrightarrow\quad & p(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sigma_1\cdots\sigma_n} \exp\Big\{-\frac{1}{2}\sum_{i=1}^n \frac{(x_i-\mu_i)^2}{\sigma_i^2}\Big\} \\
& \qquad = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\} \\
\Longrightarrow\quad & \xi_1, \cdots, \xi_n \text{ are indept.}
\end{aligned}
$$

3.2.4 Moments

3.2.4 Moments

- Origin moment of order k: $m_k = E\xi^k$.

  $E\xi = m_1$.

3.2.4 Moments

- Origin moment of order k: $m_k = E\xi^k$.
  $E\xi = m_1$.
- Center moment of order k: $c_k = E(\xi - E\xi)^k$.
  $Var\xi = c_2$.

3.2.4 Moments

- Origin moment of order k: $m_k = E\xi^k$.
  $E\xi = m_1$.
- Center moment of order k: $c_k = E(\xi - E\xi)^k$.
  $Var\xi = c_2$.
- Skewness:

$$\frac{c_3}{c_2^{3/2}}$$

3.2.4 Moments

- Origin moment of order k: $m_k = E\xi^k$.
  $E\xi = m_1$.
- Center moment of order k: $c_k = E(\xi - E\xi)^k$.
  $Var\xi = c_2$.
- Skewness:
  $$\frac{c_3}{c_2^{3/2}}$$
- kurtosis coefficient:
  $$\frac{c_4}{c_2^2} - 3$$

.

Example 10. $\xi \sim N(0, \sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n$$

Example 10. $\xi \sim N(0, \sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n = \sigma^n E\left(\frac{\xi}{\sigma}\right)^n = \sigma^n EN(0,1)^n$$

Example 10. $\xi \sim N(0, \sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n = \sigma^n E\left(\frac{\xi}{\sigma}\right)^n = \sigma^n EN(0,1)^n$$

$$= \sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx$$

Example 10. $\xi \sim N(0, \sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n = \sigma^n E\left(\frac{\xi}{\sigma}\right)^n = \sigma^n EN(0,1)^n$$

$$= \sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx = -\sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n-1} de^{-\frac{x^2}{2}}$$

Example 10. $\xi \sim N(0, \sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n = \sigma^n E\left(\frac{\xi}{\sigma}\right)^n = \sigma^n EN(0,1)^n$$

$$= \sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx = -\sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n-1} de^{-\frac{x^2}{2}}$$

$$= \sigma^n \left[-\frac{1}{\sqrt{2\pi}} x^{n-1} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + (n-1) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n-2} e^{-\frac{x^2}{2}} dx \right]$$

Example 10. $\xi \sim N(0,\sigma^2)$, then $E\xi = 0$ and

$$m_n = c_n = E\xi^n = \sigma^n E\left(\frac{\xi}{\sigma}\right)^n = \sigma^n EN(0,1)^n$$

$$= \sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx = -\sigma^n \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n-1} de^{-\frac{x^2}{2}}$$

$$= \sigma^n \left[-\frac{1}{\sqrt{2\pi}} x^{n-1} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + (n-1)\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{n-2} e^{-\frac{x^2}{2}} dx \right]$$

$$= \begin{cases} 0, & n = 2k+1, \\ 1 \cdot 3 \cdots (n-1)\sigma^n, & n = 2k. \end{cases}$$

In particular, $c_3 = 0$, $m_4 = c_4 = 3\sigma^4$. Hence for an arbitrary σ, the normal distribution has 0 skewness and kurtosis.

We can use the origin moments to express the center moments:

$$c_k = \sum_{r=0}^{k} (-1)^r \binom{k}{r} m_1^r m_{k-r}.$$

Conversely, we can also use center moments to express origin moments:

$$m_k = \sum_{r=0}^{k} (-1)^r \binom{k}{r} m_1^r c_{k-r}.$$

- **Absolute moment** of order α: $M_\alpha = E|\xi|^\alpha$, where α is a real number.

- Absolute moment of order α: $M_\alpha = E|\xi|^\alpha$, where α is a real number.

$\xi \sim N(0, \sigma^2)$:

- Absolute moment of order α: $M_\alpha = E|\xi|^\alpha$, where α is a real number.

$\xi \sim N(0, \sigma^2)$:

$$E|\xi|^n = \begin{cases} \sqrt{\frac{2}{\pi}} 2^k k! \sigma^{2k+1}, & n = 2k+1, \\ 1 \cdot 3 \cdots (n-1)\sigma^n, & n = 2k. \end{cases}$$

In general, for $\xi \sim N(0, \sigma^2)$ and $\alpha > 0$,

$$E|\xi|^\alpha = \sigma^\alpha E|N(0,1)|^\alpha$$

In general, for $\xi \sim N(0,\sigma^2)$ and $\alpha > 0$,

$$\begin{aligned} E|\xi|^\alpha &= \sigma^\alpha E|N(0,1)|^\alpha \\ &= \sigma^\alpha \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^\alpha e^{-\frac{x^2}{2}} dx \end{aligned}$$

In general, for $\xi \sim N(0, \sigma^2)$ and $\alpha > 0$,

$$\begin{aligned} E|\xi|^\alpha &= \sigma^\alpha E|N(0,1)|^\alpha \\ &= \sigma^\alpha \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^\alpha e^{-\frac{x^2}{2}} dx \\ &= \sigma^\alpha \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^\alpha e^{-\frac{x^2}{2}} dx \end{aligned}$$

In general, for $\xi \sim N(0, \sigma^2)$ and $\alpha > 0$,

$$\begin{aligned}
E|\xi|^\alpha &= \sigma^\alpha E|N(0,1)|^\alpha \\
&= \sigma^\alpha \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^\alpha e^{-\frac{x^2}{2}} dx \\
&= \sigma^\alpha \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^\alpha e^{-\frac{x^2}{2}} dx \\
&\overset{y=x^2/2}{=} \sigma^\alpha \sqrt{\frac{2}{\pi}} 2^{\frac{\alpha-1}{2}} \int_0^{\infty} y^{\frac{\alpha+1}{2}-1} e^{-y} dy
\end{aligned}$$

In general, for $\xi \sim N(0,\sigma^2)$ and $\alpha > 0$,

$$
\begin{aligned}
E|\xi|^\alpha &= \sigma^\alpha E|N(0,1)|^\alpha \\
&= \sigma^\alpha \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x|^\alpha e^{-\frac{x^2}{2}} dx \\
&= \sigma^\alpha \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} x^\alpha e^{-\frac{x^2}{2}} dx \\
&\overset{y=x^2/2}{=} \sigma^\alpha \sqrt{\frac{2}{\pi}} 2^{\frac{\alpha-1}{2}} \int_0^\infty y^{\frac{\alpha+1}{2}-1} e^{-y} dy \\
&= \sigma^\alpha \sqrt{\frac{2}{\pi}} 2^{\frac{\alpha-1}{2}} \Gamma\left(\frac{\alpha+1}{2}\right).
\end{aligned}
$$

Example 11. If $\xi \sim E(\lambda)$, then we have for $k \geq 1$,

$$E\xi^k =$$

Example 11. If $\xi \sim E(\lambda)$, then we have for $k \geq 1$,

$$E\xi^k = \int_0^\infty x^k \lambda e^{-\lambda x} dx =$$

Example 11. If $\xi \sim E(\lambda)$, then we have for $k \geq 1$,

$$E\xi^k = \int_0^\infty x^k \lambda e^{-\lambda x} dx = \frac{k}{\lambda} E\xi^{k-1},$$

Example 11. If $\xi \sim E(\lambda)$, then we have for $k \geq 1$,

$$E\xi^k = \int_0^\infty x^k \lambda e^{-\lambda x} dx = \frac{k}{\lambda} E\xi^{k-1},$$

from which we further recursively derive

$$E\xi^k = \frac{k!}{\lambda^k}.$$

Moment generating functions

For ξ, people usually define its moment generating function by

$$M_\xi(t) = Ee^{t\xi} = \int_{-\infty}^{\infty} e^{tx} dF_\xi(x), \quad t \in T$$

for some $T \subseteq \mathbf{R}$ provided that the required expected values exist.

Example. If $\xi \sim N(\mu, \sigma^2)$, then

$$M_\xi(t) = Ee^{t\xi} = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, \qquad t \in \mathbf{R}.$$

Example. If $\xi \sim N(\mu, \sigma^2)$, then

$$M_\xi(t) = Ee^{t\xi} = e^{\mu t + \frac{1}{2}\sigma^2 t^2}, \qquad t \in \mathbf{R}.$$

Example. If $\xi \sim E(\lambda)$, then when $t < \lambda$,

$$M_\xi(t) = \frac{\lambda}{\lambda - t};$$

when $t \geq \lambda$, $M_\xi(t)$ does not exist.

1. If ξ has *mgf* $M_\xi(t)$, then

$$E\xi^n = \frac{d^n}{dt^n} M_\xi(t)\Big|_{t=0};$$

1. If ξ has *mgf* $M_\xi(t)$, then
$$E\xi^n = \frac{d^n}{dt^n} M_\xi(t)\Big|_{t=0};$$
2. Suppose ξ has *mgf* $M_\xi(t)$ and η has *mgf* $M_\eta(t)$. If $M_\xi(t) = M_\eta(t)$ for all t in some neighborhood of 0, then $F_\xi = F_\eta$;

1. If ξ has *mgf* $M_\xi(t)$, then

$$E\xi^n = \frac{d^n}{dt^n} M_\xi(t)\Big|_{t=0};$$

2. Suppose ξ has *mgf* $M_\xi(t)$ and η has *mgf* $M_\eta(t)$. If $M_\xi(t) = M_\eta(t)$ for all t in some neighborhood of 0, then $F_\xi = F_\eta$;
3. If ξ has *mgf* $M_\xi(t)$, $t \in T_1$, η has *mgf* $M_\eta(t)$, $t \in T_2$, and ξ, η are independent, then

$$M_{\xi+\eta}(t) = M_\xi(t) M_\eta(t), \quad t \in T_1 \cap T_2.$$

Moment generating function is an important tool in the study of random variables and distribution functions, but it does not necessarily exist for all t.