WEEK 7

Week 7 theme

Determine the expression of ordinary moments of order k for a random variable that follows:

a) Law Gamma
b) Beta Law
c) Law χ^2 (hi-squared)
d) Legea Student
e) Legea Fisher-Snedecor

Moments

- It is called **the initial (regular) moment of order k** of random variable X, the numerical characteristic $\sigma_k = M(X^k)$
- For k=1 we have $\sigma_1 = M(X)$ and for k=2, $D^2(X) = \sigma_2 - \sigma_1^2$
- If X is a discrete type variable with the distribution $X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i \in I}$, $\sum_{i \in I} p_i = 1$ then $\sigma_k = \sum_{i \in I} x_i^k p_i$
- If X is variable of continuous type $X : \begin{pmatrix} x \\ \rho(x) \end{pmatrix}_{x \in R}$, then $\sigma_k = \int_R x^k \rho(x) dx$
- It is called **the centered moment of order k** of random variable X, the numerical characteristic $\mu_k = M\left[(X - M(X))^k\right]$, i.e.

$$\mu_k = \begin{cases} \sum_{i \in I} (x_i - M(X))^k \cdot p_i & \text{, X discretă} \\ \int_R (x - M(X))^k \rho(x) dx & \text{, X continuă} \end{cases}$$

- For k=1 we have $\mu_1 = 0$, and for k=2, $\mu_2 = D^2(X)$
- **Theorem** Between centered and initial moments there is the following relationship: $\mu_k = \sum_{i=0}^{k} (-1)^i C_k^i \sigma_{k-i} \sigma_1^i$.
- In mathematical statistics, the first four centered moments are usually used: $\mu_1, \mu_2, \mu_3, \mu_4$.

- The **initial moment of order (r,s)** of the random vector (X,Y) is called the numerical characteristic $\sigma_{rs} = M(X^r Y^s)$, i.e.

$$\sigma_{rs} = \begin{cases} \sum_{i\in I}\sum_{j\in J} x_i^r y_j^s p_{ij} & ,(X,Y) \text{ discret} \\ \iint_{R^2} x^r y^s \rho(x,y)dxdy & ,(X,Y) \text{ continuu} \end{cases}$$

- It is called **the order centered momentum (r,s)** of the random vector (X,Y), the numerical characteristic $\mu_{rs} = M\left[(X-M(X))^r (Y-M(Y))^s\right]$, i.e.

$$\mu_{rs} = \begin{cases} \sum_{i\in I}\sum_{j\in J} (x_i - M(X))^r (y_j - M(Y))^s p_{ij} & ,(X,Y) \text{ discret} \\ \iint_{R^2} (x-M(X))^r (y-M(Y))^s \rho(x,y)dxdy & ,(X,Y) \text{ continuu} \end{cases}$$

Observation

$\sigma_{10} = M(X), \sigma_{01} = M(Y), \mu_{20} = D^2(X), \mu_{02} = D^2(Y)$

Correlation or covariance

- It's called **the correlation or covariance** of random variables X and Y, the numerical characteristic

$C(X,Y) = M[(X-M(X))(Y-M(Y))]$ adică $C(X,Y) = \mu_{11}$

Observation

$C(X,Y) = M(XY) - M(X)M(Y)$, $C(X,Y) = \sigma_{11} - \sigma_{10}\sigma_{01}$

If X, Y independent $\Rightarrow C(X,Y) = 0$ but not reciprocal.

$C(X,X) = D^2(X)$

$C\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j C(X_i, Y_j)$, whatever the random variables Xi and Yj and whatever the real constants ai and b_j, $1 \le i \le m, 1 \le j \le n$

$C(X,Y) = C(Y,X)$, whatever X and Y.

- It is called **the correlation coefficient** relative to random variables X and Y numerical characteristic

$$r(X, Y) = \frac{C(X,Y)}{\sqrt{D^2(X)}\sqrt{D^2(Y)}}$$

Observation

1) X, Y $\Rightarrow r(X,Y) = 0$ independently of each other is not true;

2) We say that X,Y are uncorrelated if r(X,Y) =0

Properties:

a) $|r(X,Y)| \le 1$

b) $r(X,Y) = +1 \Leftrightarrow Y = a\,\mathrm{X} + \mathrm{b}, a > 0$

c) $r(X,Y) = -1 \Leftrightarrow Y = aX + b, a < 0$

- In practice it is also said that:

1) X and Y are positively perfectly correlated if $r(X,Y) = 1$;

2) X and Y are negatively correlated perfectly if $r(X,Y) = -1$;

3) X and Y are strongly positively (or negatively) correlated if $0{,}75 \le r(X,Y) < 1$ (or $-1 < r(X,Y) \le -0{,}75$);

4) X and Y are weakly positive (or negative) correlated if $0 < r(X,Y) < 0{,}25$ (or $-0{,}25 < r(X,Y) < 0$);

The decision-making value margins are conventionally chosen.

- Given the random vector $Z = (X_1, X_2, ..., X_n)$ $Z : E \to R^n$, it is called **its mean value** and denoted by $M(Z)$, if it exists, the n-dimensional vector whose components are the average values of the components of Z, i.e.: $M(Z) = M(X_1, X_2, ..., X_n) = (M(X_1), M(X_2), ..., M(X_n))$.

- It is called **the covariance** (or correlation) matrix of the vector Z, and is denoted by $C(Z)$, if any, the matrix $C(Z) = (c_{ij})_{\substack{i=\overline{1,n} \\ j=\overline{1,n}}} = (C(X_i, Y_j))_{i,j=\overline{1,n}}$

Observation

a) For the case of a two-dimensional random vector, do not confuse the mean of the product of components X and Y, which is $M(XY)$ with the mean of the vector (X,Y), which is $M(X,Y)$.

b) Sometimes the correlation matrix $C(Z)$ is also denoted by $\Gamma(Z)$.

c) Unfolded, the covariance matrix $C(Z)$ has the form:

$$C(Z) = \begin{pmatrix} D^2(X_1) & C(X_1, X_2) & \dots & C(X_1, X_n) \\ C(X_2, X_1) & D^2(X_2) & \dots & C(X_2, X_n) \\ \dots & \dots & \dots & \dots \\ C(X_n, X_1) & C(X_n, X_2) & \dots & D^2(X_n) \end{pmatrix}$$

And as a result of the properties of correlation, we find that the matrix $C(Z)$ is symmetric.

d) Starting from the definition of correlation coefficient and correlation matrix, if all components of Z are non-constant, then we can introduce the matrix of correlation coefficients $R(Z)$ whose developed form is:

$$R(Z)=\begin{pmatrix} 1 & r(X_1,X_2) & \dots & r(X_1,X_n) \\ r(X_2,X_1) & 1 & \dots & r(X_2,X_n) \\ \dots & \dots & \dots & \dots \\ r(X_n,X_1) & r(X_n,X_2) & \dots & 1 \end{pmatrix}$$

Both forms of the correlation matrix of the random vector Z are actually tables measuring the degree of dependence between the components of Z, considered two by two.

Application 1. *Let (X,Y) be a discrete random vector whose probabilistic distribution is given by the table below.*

Calculate the correlation coefficient r(X,Y).

And / X	-1	0	1	2	pi
-1	1/6	1/12	1/12	1/24	9/24
0	1/24	1/6	1/12	1/24	8/24
1	1/24	1/24	1/6	1/24	7/24
qj	6/24	7/24	8/24	3/24	1

Solution:

Based on the appropriate formulas, we immediately deduce:

$$M(X)=-1\frac{9}{24}+0\frac{8}{24}+1\frac{7}{24}=-\frac{2}{24}=-\frac{1}{12}$$

$$M(X^2)=1\frac{9}{24}+0\frac{8}{24}+1\frac{7}{24}=\frac{16}{24}=\frac{2}{3}$$

$$D(X^2)=M(X^2)-[M(X)]^2=\frac{2}{3}-\frac{1}{144}=\frac{95}{144}$$

$$M(Y)=-1\frac{6}{24}+0\frac{7}{24}+1\frac{8}{24}+2\frac{3}{24}=\frac{8}{24}=\frac{1}{3}$$

$$M(Y^2)=1\frac{6}{24}+0\frac{7}{24}+1\frac{1}{24}+4\frac{3}{24}=\frac{26}{24}=\frac{13}{12}$$

$$D(Y^2)=M(Y^2)-[M(Y)]^2=\frac{13}{12}-\frac{1}{9}=\frac{35}{36}$$

$$M(XY)=-1\cdot\left(-1\frac{1}{6}+0\frac{1}{12}+1\frac{1}{12}+2\frac{1}{24}\right)+0\cdot\left(-1\frac{1}{24}+0\frac{1}{6}+1\frac{1}{12}+2\frac{1}{24}\right)+$$
$$+1\cdot\left(-1\frac{1}{24}+0\frac{1}{24}+1\frac{1}{6}+2\frac{1}{24}\right)=\frac{5}{24}$$

$$C(X,Y)=M(XY)-M(X)M(Y)=\frac{5}{24}+\frac{1}{36}=\frac{17}{72}$$

$$r(X,Y)=\frac{C(X,Y)}{\sqrt{D^2(X)D^2(Y)}}=\frac{\frac{17}{12}}{\sqrt{\frac{95}{144}\cdot\frac{35}{36}}}\approx 0{,}295$$

Application 2. *Let be random variables:*

$X_1:\begin{pmatrix}-1 & 1\\ p_1 & p_2\end{pmatrix}$; $X_2:\begin{pmatrix}1 & 3\\ q_1 & q_2\end{pmatrix}$; $X_3:\begin{pmatrix}2 & 4\\ r_1 & r_2\end{pmatrix}$

whose common distribution denoted (p_{ijk}), $1\le i,j,k\le 2$, *is:*

$p_{111}=\frac{1}{16}$; $p_{112}=\frac{1}{16}$; $p_{121}=\frac{1}{32}$; $p_{122}=\frac{3}{32}$;

$p_{211}=\frac{1}{8}$; $p_{212}=\frac{1}{4}$; $p_{221}=\frac{1}{16}$; $p_{222}=\frac{5}{16}$.

Determine the two-dimensional and one-dimensional distributions of the three-dimensional random vector $Z=(X_1,X_2,X_3)$ *and correlation matrices* $C(Z)$ *and* $R(Z)$.

Solution:

We immediately have two-dimensional distributions

$$\begin{cases}p_{11\bullet}=p_{111}+p_{112}=\frac{1}{8}\\ p_{21\bullet}=p_{211}+p_{212}=\frac{3}{8}\end{cases};\quad \left.\begin{matrix}p_{12\bullet}=p_{121}+p_{122}=\frac{1}{8}\\ p_{22\bullet}=p_{221}+p_{222}=\frac{3}{8}\end{matrix}\right\}\text{ for }(X_1,X_2)$$

$$\begin{cases}p_{1\bullet 1}=p_{111}+p_{121}=\frac{3}{32}\\ p_{2\bullet 1}=p_{211}+p_{221}=\frac{8}{16}\end{cases};\quad \left.\begin{matrix}p_{1\bullet 2}=p_{112}+p_{122}=\frac{5}{32}\\ p_{2\bullet 2}=p_{212}+p_{222}=\frac{9}{16}\end{matrix}\right\}\text{ for }(X_1,X_3)$$

$$\begin{cases}p_{\bullet 11}=p_{111}+p_{211}=\frac{3}{16}\\ p_{\bullet 21}=p_{121}+p_{221}=\frac{3}{32}\end{cases};\quad \left.\begin{matrix}p_{\bullet 12}=p_{112}+p_{212}=\frac{5}{32}\\ p_{\bullet 22}=p_{122}+p_{222}=\frac{13}{32}\end{matrix}\right\}\text{ for }(X_2,X_3)$$

And as a result we can write the following two-dimensional distribution tables:

X2 / X1	1	3	pi
-1	1/8	1/8	1/4
1	3/8	3/8	¾
qj	1/2	1/2	1

X3 / X1	2	4	pi
-1	3/32	5/32	1/4
1	3/16	9/16	3/4
rk	9/32	23/32	1

X3 / X2	2	4	qi
1	3/16	5/16	1/2
3	3/32	13/32	1/2
rk	9/32	21/32	1

from which are also observed one-dimensional distributions (distributions of random variables considered X1, X2, X3). From these tables we deduce by immediate calculation:

$M(X_1)=\frac{1}{2}$; $M(X_1^2)=1$; $D^2(X_1)=\frac{3}{4}$

$M(X_2)=2$; $M(X_2^2)=5$; $D^2(X_2)=1$

$M(X_3)=\frac{55}{16}$; $M(X_3^2)=\frac{101}{8}$; $D^2(X_3)=\frac{207}{256}$

$M(X_1\cdot X_2)=1$; $M(X_1)M(X_2)=1$; $C(X_1,X_2)=0$; $r(X_1,X_2)=0$

$M(X_1\cdot X_3)=\frac{29}{16}$; $M(X_1)M(X_3)=\frac{55}{32}$; $C(X_1,X_3)=\frac{3}{32}$; $r(X_1,X_3)=\sqrt{\frac{3}{207}}\approx 0{,}12$

$M(X_2\cdot X_3)=\frac{113}{16}$; $M(X_2)M(X_3)=\frac{55}{8}$;

$C(X_2,X_3)=\frac{3}{16}$; $r(X_2,X_3)=\frac{3}{\sqrt{207}}\approx 0{,}21$

And as a result we can write correlation matrices:

$$C(Z)=\begin{pmatrix} 3/4 & 0 & 3/32 \\ 0 & 1 & 3/16 \\ 3/32 & 3/16 & 207/256 \end{pmatrix} \text{ And } R(Z)=\begin{pmatrix} 1 & 0 & 0{,}12 \\ 0 & 1 & 0{,}21 \\ 0{,}12 & 0{,}21 & 1 \end{pmatrix}$$

finding that X1 and X2 are independent, while between X3 and X1 or X3 and X2 there is some dependence, even if it is not strong.

Other numerical characteristics

- It is called **the median** of a random variable X, the numerical characteristic Me that verifies the relationship: $P(X\geq M_e)\geq\frac{1}{2}\leq P(X\leq M_e)$
- If F is the distribution function and is continuous M_e, then it is determined from the equation F(M,e$=\frac{1}{2}$).
- If $M_e\in(a,b]$ then it is taken $M_e=\frac{a+b}{2}$
- A **modal value or modulus** of the random variable X is called any local maximum point of the distribution of X (in the discrete case) and probability density (in the continuous case).
- If there is only one local maximum, we say that X's law is unimodal, otherwise we call it plurimodal.
- The **asymmetry** (Fischer's coefficient) of the random variable X is called the numerical characteristic defined by $s=\frac{\mu_3}{\sigma^3}$.

- It is called **an excess** of random variable X, the numerical characteristic defined by $e = \frac{\mu_4}{\sigma^4} - 3$

.

- If e<0 then the distribution graph has a flattened appearance and the law is called platicurtic.
- If e>0 then the distribution graph has a sharp aspect and the law will be called leptocurtic.
- If e = 0 then the distributions are mesocurtic.
- If X is a random variable with the distribution function $F(x)$, the **quartiles (in number of three) of X (or of the distribution of X)** are called the numbers q_1, q_2 and q_3 with the properties:

$$\begin{cases} F(q_1) \le \frac{1}{4} \\ F(q_1 + 0) \ge \frac{1}{4} \end{cases} \qquad \begin{cases} F(q_2) \le \frac{1}{2} \\ F(q_2 + 0) \ge \frac{1}{2} \end{cases} \qquad \begin{cases} F(q_3) \le \frac{3}{4} \\ F(q_3 + 0) \ge \frac{3}{4} \end{cases}$$

We note that $q_2 = M_e$.