ERASMUS, B4 Course
Winter 2020/2021
FPM&St., December 3

Subject. Markov & Tschebyscheff Inequality – the stochastic interpretation of the MEAN VALUE.

We will start by summarizing the results obtained.

The central point of our investigation is a RANDOM Experiment and its observation process, which consists of:

(i) KPM – (Ω, Σ, P)

(ii) the model interpreter X – random variable

As a result of combining these elements we get the probability distribution, whose basic description is the notion of the cummulative probability function, F_X, where

$$\mathbb{R} \ni t \longrightarrow P(\{\omega \in \Omega : X(\omega) < t\}) = F_X(t)$$

By theoretical reasons we have reduced our investigation to two cases of X:

a) X discrete, then F_X is a step function interpreted by the distribution d_X,

$$d_X \quad \begin{array}{c|c|c|c} x_1 & x_2 & \ldots & x_n \\ \hline p_1 & p_2 & \ldots & p_n \end{array}$$

$X(\Omega) = \{x_1, \ldots, x_n\}$

$\forall_{1 \le k \le n} \quad p_k = P(\{\omega \in \Omega : X(\omega) = x_k\})$ (for finite case)

b) continuous: then F_X is a smooth function, so there exists F_X' and

$f_X \stackrel{df}{=} F_X'$ is called a density prob. function, where

$$P(\{\omega \in \Omega : X(\omega) < t\}) = F_X(t) = \int_{-\infty}^{t} f_X(u)\,du$$

For given X we have defined two parameters:

$m_X = EX$ - the mean value (or math. expectation),

$\sigma_X^2 = var(X)$ - the variance.

From the mathematical point of view, $EX = \int_\Omega X\,dP$, and we know that we have respectively

$$EX = \sum x_k p_k \quad , \quad EX = \int_R t f_X(t)\,dt .$$

Moreover, from the existence of $E(X^2)$ follows the existence of EX, and $var(X) = m_2 - m_1^2$.

PROBLEM: WHY m_X and σ_X^2 are so important in P.T.?

We begin starting with two remarks. ~~As we tr~~

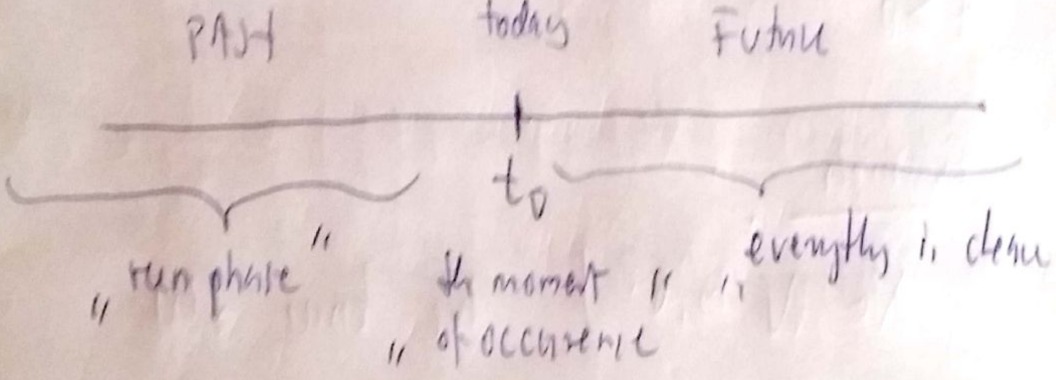
1°. As we know, the main goal of the P.T. is to give the whole description of given R.V.

In discrete case we have the following situation:

$X(\Omega) = \{x_1, x_2, \dots, x_n\}$

$P(\{\omega \in \Omega : X(\omega) = x_i\}) = p_i$

Look at the observing process as a function of time



• Until the moment t_0 we do not know what will happen surely. We only know that

• In the moment t_0 and after that, we know everything: only one $x_j \in X(\Omega)$ is a result of the observation

From the point of view of application of P.T., the evaluation of the observation result should take place for the moment $t < t_0$ at the least, because for at time t_0 ~~should~~ a decision must be made which influences behavior of the system when our RE exists.

How then should this data

x_1	x_2	$\dots$	x_n
p_1	p_2	$\dots$	p_n

be understood?

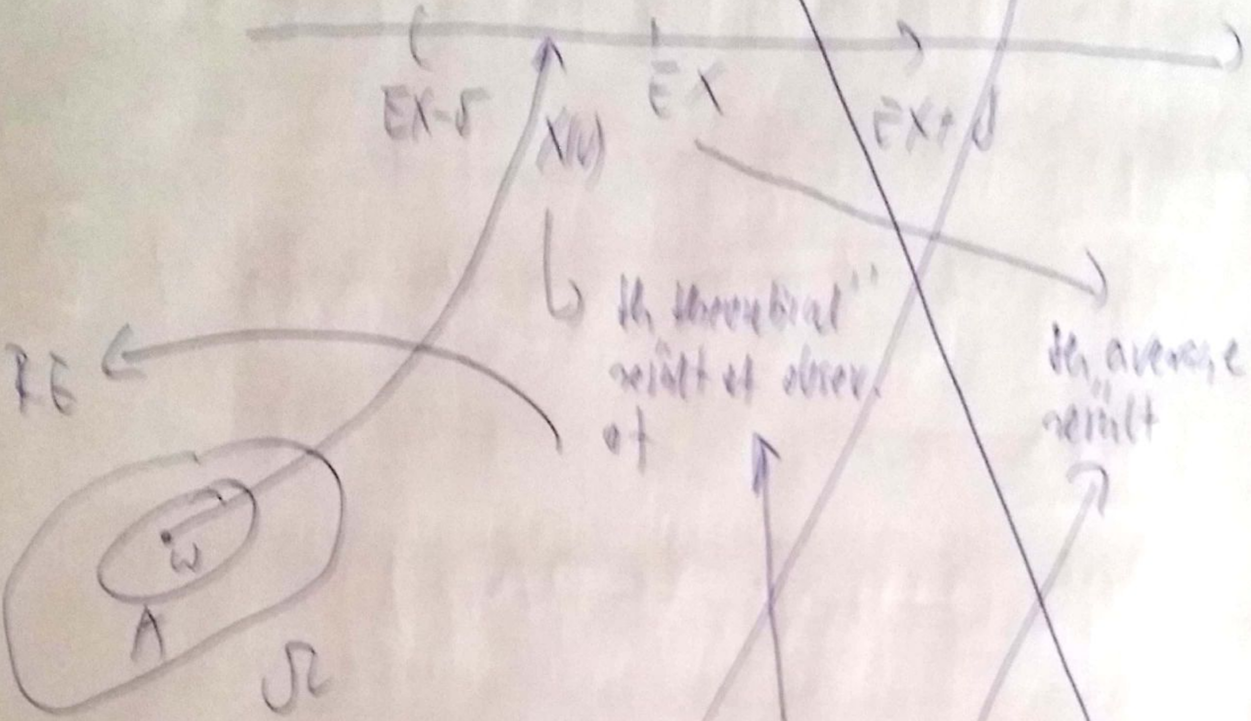
We need some ~~a~~ selection criterion!

In the continuous case the situation is more complicated, because, as we know then

$$\forall_{t_0 \in \mathbb{R}} \; P(\{v \in \Omega : X(v) = t_0\}) = 0 \; (!!!)$$

We are going to show, the solution of the problem is m_X and σ_X^2 (whenever exists); namely: ~~a~~ we will show that if X has a second moment, that the result of one observation is $m_X = EX$, where

(*) $\forall_{\varepsilon>0} \; \exists_{\delta>0} \; P(\{\omega \in \Omega : |X(\omega) - EX| < \delta\}) \geq 1 - \varepsilon$,



[So in (*) we ~~connect~~ connect X(ω) to EX taking into account the stochastic control of the dispersion effect

To prove (*) we need so called Tschebyscheff inequality.

First we prove the simpler version of T.N. – the Markov .

Th 1 (the Markov inequality)

Assumptions: $X(\Omega) \subset \mathbb{R}_+ = [0, +\infty)$ and EX exists.

Then

$$\forall_{a>0} \quad (**) \quad P(\{\omega \in \Omega : X(\omega) \geq a\}) \leq \frac{1}{a} EX \quad (\text{Markov}).$$

Proof (discrete case)

By assumption $X(\Omega) = \{a_n \in \mathbb{R}_+, \, n \in \mathbb{N}_0 \, (\subset \mathbb{N})\}$,

$P(\{\omega \in \Omega : X(\omega) = a_n\}) = p_n$.

We have successively for $a > 0$:

$$EX = \sum_{n \in \mathbb{N}_0} a_n p_n \geq \sum_{\substack{n \in \mathbb{N}_0 \\ a_n \geq a}} a_n p_n \geq \sum_{\substack{n \in \mathbb{N}_0 \\ a_n \geq a}} a \, p_n =$$

$$= a \sum_{\substack{n \in \mathbb{N}_0 \\ a_n \geq a}} p_n = a \, P(\{\omega \in \Omega : X(\omega) \geq a\}).$$

For continuous case we can write

$$EX = \int_{\mathbb{R}} t f_X(t) dt \geq \int_{\{t : t \geq a\}} t f_X(t) dt \geq \int_{\{t : t \geq a\}} a f_X(t) dt$$

$$= a \int_{\{t:\, t \geq a\}} f_X(t)\,dt = a\,P(\{\omega \in \Omega : X(\omega) \geq a\})$$

//.

Examples.

1) $X \in E(\lambda)$, $\lambda > 0$ and $a = \lambda$

Then from Markov

$$P(\{\omega \in \Omega : X(\omega) \geq \lambda\}) \leq \frac{1}{\lambda} EX = \frac{1}{\lambda^2}$$

TASK 1. Give an geometrical interpretation of the above result.

2) $X \in B(n, p)$. Then for every $k \in \{1, 2, \dots, n\}$

$$P(\{\omega \in \Omega : X(\omega) \geq k\}) \leq \frac{1}{k} np$$

So, let $k = 50\%$ of successes, so $k = n/2$

then

$$P(\{\omega \in \Omega : X(\omega) \geq n/2\}) \leq 2p$$

TASK 2.

Show the application of Markov inequality for gaussian distribution.

Now we are ready to study the Tshebysheff ineq.

Assume that X has at least 2-nd moment, and $X \neq$ const.

We define (for a moment) a new random var.

$$Y \stackrel{def}{=} (X-a)^2, \text{ where } a \in \mathbb{R} \text{ is fixed.}$$

We note that for Y we can apply the Markov ineq.

So, $\forall t > 0$, from $(**)$ we get

$$\boxed{P(\{\omega \in \Omega : Y(\omega) \geq t^2\}) \leq \frac{1}{t^2} EY, \quad (t>0)}$$

But def. of Y,

$$P(\{\omega \in \Omega : (X(\omega)-a)^2 \geq t^2\}) \leq \frac{1}{t^2} E(X-a)^2,$$

or equivalently

$$P(\{\omega \in \Omega : |X(\omega)-a| \geq t\}) \leq \frac{1}{t^2} E(X-a)^2.$$

Finally, let $a = m_X = EX$, and we substitute $t \to t\sigma$, $\sigma = \sigma_X = \sqrt{\sigma_X^2}$

Then we obtain

$$P(\{v \in \Omega : |X(v) - m| \geq t\sigma\}) \leq \frac{1}{t^2\sigma^2} E(X-m)^2 = \frac{1}{t^2}$$

Therefore we have

<u>Th 2 (Tschebyscheff)</u>

If X = const and has its 2-nd moment, then for $m = EX$, $\sigma^2 = var(X)$ we have

(***) $\forall_{t>0} \; P(\{v \in \Omega : |X(v) - m| \geq t\sigma\}) \leq 1/t^2$

<u>Example</u>.

$X \in N(m, \sigma^2)$. We know that

$N = \frac{X-m}{\sigma} \in N(0,1)$.

So from (***) we have

$P(\{v \in \Omega : |N(v)| \geq t\}) \leq \frac{1}{t^2}$

If we fix $t = 3$, we obtain so <u>called</u>

„3σ rule" for ~~[illegible]~~ gaussian distr.

$$P(\{\omega\in\Omega : |X(\omega)-m| \geq 3\sigma\}) =$$

$$= P(\{\omega\in\Omega : |N(\omega)| \geq 3\}) \leq \frac{1}{9} \approx 0{,}111.$$

TASK 4

Estimate the above by using the notion of $\overline{\Phi}$.

TASK 5

Apply (***) for $X \in B(n,p)$

TASK 6

Apply (***) for $X \in W(\lambda)$.