O matematičkom očekivanju i disperziji

Do sada smo se susreli sa više definicija matematičkog očekivanja, zavisno od tipa slučajne veličine. Opisaćemo sada kroz primer motivaciju koja stoji iza definicije očekivanja u diskretnom slučaju.

Bacamo novčić i želimo da damo prognozu o sledećem ishodu ovog eksperimenta.
Šta bismo učinili da do sada nismo nijednom bacili novčić, tj. da nismo dobili informaciju o tome koliko je verovatan ishod pismo, a koliko glava?
Najverovatnije bismo rekli da su šanse pola-pola i da ne možemo dati prednost nekom od ova dva ishoda.
Dakle, naša predikcija se nalazi između glave i pisma i to tačno na sredini.

Matematički zapisano, slučajna veličina koja je ishodu pismo dodelila 1, a ishodu glava 0 ima sledeći zakon raspodele:

$$X : \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Definicija matematičkog očekivanja koju znamo kaže da je:

$$E(X) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$$

i zaista, $\frac{1}{2}$ jeste jednako udaljena od vrednosti 0 i 1, kojima smo označili moguće ishode.

Međutim, $\frac{1}{2}$ ne odgovara nijednom ishodu eksperimenta, pa **NE MOŽEMO** reći da je matematičko očekivanje alat za prognoziranje **TAČNIH** vrednosti slučajne veličine X, već da ono više govori o položaju te prognoze, uzimajući u obzir sve verovatnoće ishoda.

Šta bismo zaključili da smo novčić bacili 1000 puta i zabeležili sledeće rezultate?

	pismo	glava
pojavljivanja	600	400
relativna frekvencija	0.6	0.4

Sada naša prognoza više naginje ka pismu, ali ne previše. Ako definišemo verovatnoće pojavljivanja pisma i glave kao relativne frekvencije, slučajna veličina X sada ima sledeći zakon raspodele:

$$X : \begin{pmatrix} 0 & 1 \\ 0.4 & 0.6 \end{pmatrix}$$

i njeno očekivanje je:

$$E(X) = 0 \cdot 0.4 + 1 \cdot 0.6 = 0.6$$

Kako je 0.6 za nijansu bliže 1 (pismo), $E(X)$ je ponovo na odgovarajući način opisalo našu intuiciju.

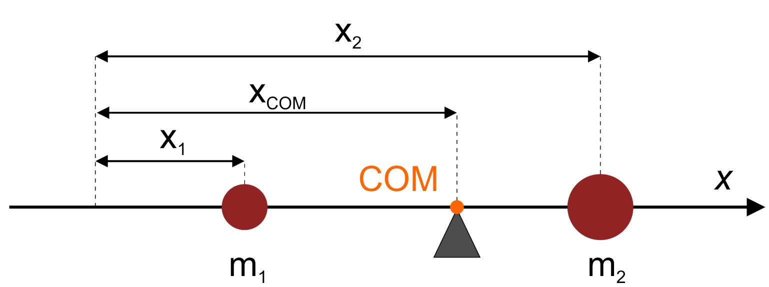
Još jedan lep primer koji može objasniti poreklo definicije matematičkog očekivanja dolazi iz fizike.

Ponovimo sada N puta eksperiment opisan slučajnom veličinom X:

$$X : \begin{pmatrix} x_1 & x_2 \\ p_1 & p_2 \end{pmatrix}$$

i za svaki ishod označen sa x_1 stavimo teg jedinične mase na tačku x_1 realne ose. Isto to radimo i za ishod označen sa x_2. Šta će se dogoditi? Ukupna masa tegova koji se nalaze na x_1 iznosiće oko $m_1 = Np_1$ jedinica, a ukupna masa na x_2 biće oko $m_2 = Np_2$ jedinica. Uz malo fizike možemo pokazati da tačka koja je centar mase ovog sistema ima koordinatu:

$$\frac{x_1 m_1 + x_2 m_2}{m_1 + m_2} = \frac{x_1 N p_1 + x_2 N p_2}{N p_1 + N p_2} = x_1 p_1 + x_2 p_2$$



Ako ispod tegova postavimo i homogenu dasku, tačka sa koordinatom $x_1p_1+x_2p_2$ je i centar ravnoteže, tj. mesto na kome možemo postaviti kamen tako da daska sa sve tegovima stoji u horizontalnom položaju.

A kako je i očekivanje baš definisano sa $E(X) = x_1p_1 + x_2p_2$, vidimo da $E(X)$ možemo shvatiti i kao tačku koja drži u ravnoteži sve moguće vrednosti slučajne veličine X, pri čemu je relativna težina svake tačke x_i određena verovatnoćom p_i.

Kako teže tačke vuku centar mase ka sebi, $E(X)$ je bliže onim tačkama čija je verovatnoća realizacije veća, što se u potpunosti slaže s našom intuicijom.

Šta je disperzija?

Rekli smo da $E(X)$ ne daje tačnu prognozu slučajne veličine X, ali nam ipak na neki način može pomoći u prognoziranju. Ono što se neretko čini je da se prognoza slučajne veličine X da kao tačka iz skupa mogućih vrednosti X koja je najbliža $E(X)$. Posmatrajmo sada dve slučajne veličine sa istim očekivanjima, ali različitim raspodelama:

$$X_1 : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}, \quad X_2 : \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Kada bismo prognoze slučajnih veličina X_1 i X_2 dali na opisani način, u prvom slučaju bismo često grešili, jer su tačke $1, 2, 4$ i 5 podjednako zastupljene u eksperimentima kao i 3 (naša prognoza). Međutim, u drugom slučaju nikada ne bismo pogrešili, jer X_2 može uzeti samo vrednost 3. Česte greške u prvom slučaju su posledica postojanja tačaka dosta udaljenih od očekivanja, a koje nisu malo verovatne (tj. nisu male težine). Dakle, priroda slučajne veličine X nije dovoljno opisana njenim očekivanjem, te se javila potreba za definisanjem veličine koja bi na neki način izmerila njenu *raspršenost*. Prvobitna ideja je bila da ta mera raspršenosti bude $E(X - E(X))$, ali ona je uvek jednaka 0. Sledeći predlog, koji je i usvojen, bio je da ta mera glasi:

$$E(X - E(X))^2$$

Kvadrat razlike X i $E(X)$ je ovu meru, koju zovemo disperzijom, učinio simetričnom merom raspršenosti slučajne veličine X, jer je njom zabeležen uticaj teških tačaka koje odsupaju od $E(X)$ i sa leve i sa desne strane, dok je u prethodnom pokušaju njihov uticaj međusobno anuliran. Uporediti disperzije slučajnih veličina X_1 i X_2.