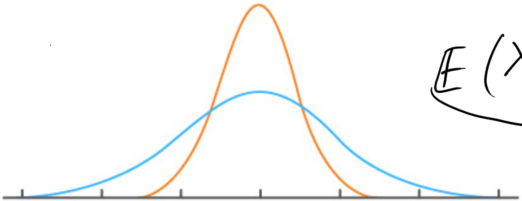
Διασπορά μιας ΤΜ.

$\mathbb{E}(X-\mu)^2$



$\mathbb{E}(X-\mu)^2$ $\mathbb{E}|X-\mu|$

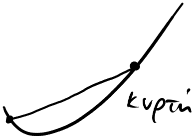
αλγεβρικά διαχειρίσιμο

Διασπορά της ΤΜ X: $\mathbb{V}ar[X] = \sigma^2(X) = \boxed{\mathbb{E}\left[(X-\mu)^2\right]}$, όπου $\mu = \mathbb{E}[X]$.

Τυπική απόκλιση της X: $\sigma(X) = \sqrt{\mathbb{V}ar[X]}$.

Διασπορά μιας ΤΜ: ένας υπολογισμός

$$
\begin{aligned}
\mathbb{V}ar[X] &= \mathbb{E}\left[(X-\mu)^2\right] \\
&= \mathbb{E}\left[X^2+\mu^2-2\mu X\right] \\
&= \mathbb{E}\left[X^2\right]+\mathbb{E}\left[\mu^2\right]-2\mu \overbrace{\mathbb{E}[X]}^{\mu} \\
&= \mathbb{E}\left[X^2\right]+\mu^2-2\mu\cdot\mu \\
\sigma^2(X) \quad &= \underbrace{\mathbb{E}\left[X^2\right]}_{2^{\eta} \text{ ροπή του } X}-(\mathbb{E}[X])^2 .
\end{aligned}
$$



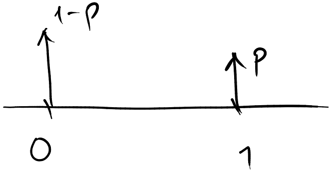
$\varphi''(x) \geq 0$

$0 \leq \mathbb{V}ar[X] = \mathbb{E}\left[X^2\right]-\mathbb{E}[X]^2$ άρα έχουμε $\boxed{\mathbb{E}[X]^2 \leq \mathbb{E}\left[X^2\right].}$

Ειδική περίπτωση ($\phi(x)=x^2$) της *ανισότητας Jensen:* $\uparrow \varphi(x) = x^2$

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)] \quad \text{για } \phi(x) \text{ κυρτή.}$$

Διασπορά μιας ΤΜ: δείκτρια



X παίρνει τις τιμές 0 ή 1, και $p = \mathbb{P}[X = 1]$.

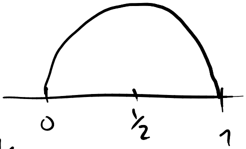
$$\mathbb{E}[X] = p \cdot 1 + (1-p) \cdot 0 = p \leftarrow$$

$$\mathbb{E}\left[X^2\right] = p \cdot 1^2 + (1-p) \cdot 0^2 = p \leftarrow$$

$$\sigma^2(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = p - p^2 = p(1-p) \leftarrow$$

↳ μεγιστ. για $p = 1/2$

$$\sigma(X) = \sqrt{p(1-p)} \leftarrow$$



Παρατήρηση: Μεγιστοποιείται για $p = \frac{1}{2}$.

Διασπορά μιας ΤΜ: διακριτή ομοιόμορφη κατανομή

X ομοιόμορφα κατανεμημένη στο $\{0, 1, 2, \ldots, m\}$. → πιθ. $\frac{1}{m+1}$

$\mathbb{E}[X] = \frac{m}{2}$ (λόγω συμμετρίας της κατανομής)

$1^2 + 2^2 + 3^2 + \cdots + m^2$

$$\mathbb{E}\left[X^2\right] = \sum_{k=0}^{m} \frac{1}{m+1} k^2 = \frac{1}{m+1} \frac{m(m+1)(2m+1)}{6} = \frac{m(2m+1)}{6}$$

$$\sigma^2(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{m(2m+1)}{6} - \frac{m^2}{4} = \frac{m(m+2)}{12} = \frac{m^2}{12} + \frac{m}{6}$$

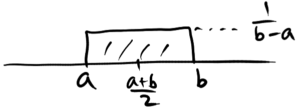
$m \to \infty$

$$\sigma(X) = \sqrt{\frac{m^2}{12} + \frac{m}{6}} = \frac{m}{\sqrt{12}} \underbrace{\sqrt{1 + \frac{2}{m}}}_{\downarrow 1 \text{ για } m \to \infty} \sim \frac{m}{\sqrt{12}} \sim (0.288\cdots) \cdot m$$

Διασπορά μιας ΤΜ: συνεχής ομοιόμορφη κατανομή

X ομοιόμορφα κατανεμημένη στο διάστημα $[a, b]$, άρα $f_X(x) = \frac{1}{b-a} : (a \leq x \leq b)$.

$$\mathbb{E}[X] = \frac{a+b}{2}$$ (λόγω συμμετρίας της κατανομής)



$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \frac{1}{b-a} \int_a^b x^2\, dx = \frac{1}{b-a} \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}$$

$$\sigma^2(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

$$\sigma(X) = \frac{(b-a)}{\sqrt{12}}$$

Διασπορά αθροίσματος ανεξαρτήτων ΤΜ

$\mu_X = \mathbb{E}X \ , \ \mu_Y = \mathbb{E}Y$

Αν X, Y ανεξάρτητες τότε $\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y)$

$$\sigma^2(X+Y) = \mathbb{E}\left[\left(X+Y-(\mu_X+\mu_Y)\right)^2\right] = \mathbb{E}\left(X-\mu_X + Y-\mu_Y\right)^2 =$$

$$= \mathbb{E}\left[(X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)(Y-\mu_Y)\right] =$$

$$= \mathbb{E}\left[(X-\mu_X)\right]^2 + \mathbb{E}\left[(Y-\mu_Y)^2\right] + 2\,\underbrace{\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]}_{=0} =$$

$$= \sigma^2(X) + \sigma^2(Y) + 2\,\cancel{\underbrace{\mathbb{E}(X-\mu_X)}_{=0}\,\underbrace{\mathbb{E}(Y-\mu_Y)}_{=0}}$$

Διασπορά αθροίσματος πολλών ανεξαρτήτων ΤΜ

Αν $X_1, X_2, \ldots, X_n$ ανεξάρτητες και $S = X_1 + X_2 + \cdots + X_n$ τότε

$$\sigma^2(S) = \sigma^2(X_1) + \cdots \sigma^2(X_n).$$

$$\sigma^2(S) = n\,\sigma^2(X_1)$$

Αν $X_1, \ldots, X_n$ ανεξάρτητες και ισόνομες τότε $\sigma(S) = \sqrt{n}\,\sigma(X_1)$.

Παρατήρηση: Αρκεί η ανεξαρτησία ανά δύο ή ακόμη και

$$\mathbb{E}\left[(X_i - \mathbb{E}[X_i]) \cdot (X_j - \mathbb{E}[X_j])\right] = 0 \text{ για } i \neq j.$$

Διασπορά διωνυμικής κατανομής

$X \in \{0, 1, \ldots, n\}$

$X \sim B(n, p)$ $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

$\mathbb{E}X = p \cdot n$

Δηλ. X μετράει επιτυχίες σε n επαναλήψεις πειράματος με πιθ. επιτυχίας p.

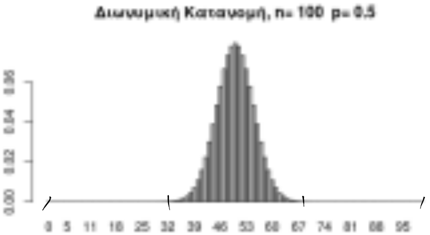
Άρα $X = X_1 + X_2 + \cdots + X_n$ όπου X_j ανεξάρτητες, ισόνομες, 0-1 ΤΜ, με $\mathbb{P}[X_1 = 1] = p$.

$\sigma^2(X) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) = n\,p(1-p)$

$\sigma(X) = \sqrt{n}\sigma(X_1) = \boxed{\sqrt{p(1-p)}}\sqrt{n}.$ $\sqrt{n} << n$

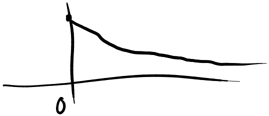
Συγκέντρωση: Εύρος τιμών n. Οι περισσότερες όμως $\sqrt{n}$ κοντά στη μέση τιμή.

Διασπορά διωνυμικής κατανομής



Διασπορά εκθετικής κατανομής

$X \sim \mathrm{Exp}(\lambda)$, δηλ. $\underline{f_X(x)} = \lambda e^{-\lambda x} : \underline{(x \geq 0)}$ (έχουμε $\underline{\underline{\lambda}} > 0$).



Έχουμε ήδη υπολογίσει $\underline{\mathbb{E}[X] = 1/\lambda}$.

$$\mathbb{E}[X^2] = \int_0^\infty x^2 \lambda e^{-\lambda x}\, dx = (\ \ldots \text{ με δις ολοκλήρωση κατά μέρη } \ldots)\ \frac{2}{\lambda^2}$$

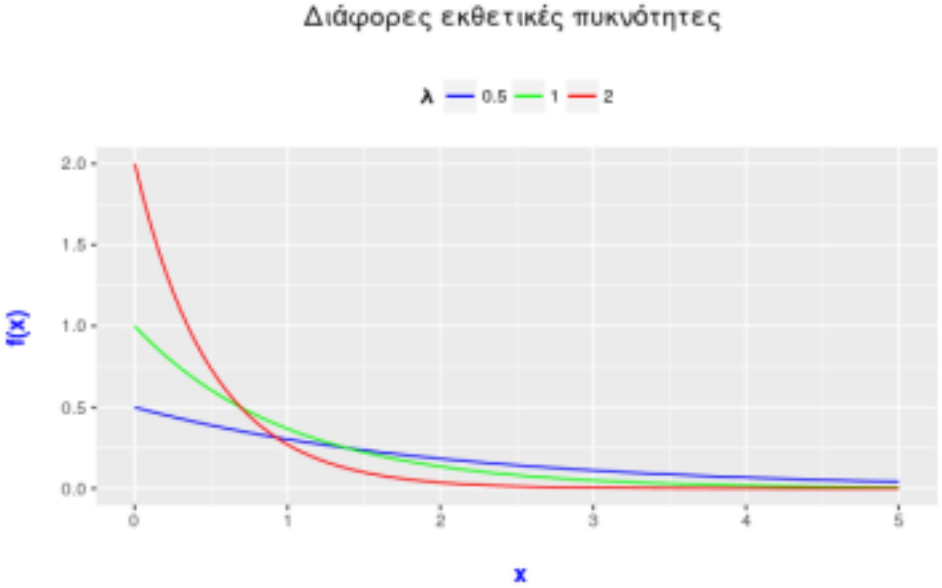
$$= \int_0^\infty x^2 (-e^{-\lambda x})'\, dx = \overset{0=}{\cancel{-x^2 e^{-\lambda x}\Big|_0^\infty}} + 2\int_0^\infty x e^{-\lambda x} dx = 2\int_0^\infty x \left(\frac{e^{-\lambda x}}{-\lambda}\right)' dx =$$

$$= \underset{0=}{\cancel{-2x\frac{e^{-\lambda x}}{\lambda}\Big|_0^\infty}} + 2\frac{1}{\lambda}\int_0^\infty e^{-\overbrace{\lambda x}^{y}} dx = 2\frac{1}{\lambda}\int_0^\infty e^{-y}\frac{dy}{\lambda} = \frac{2}{\lambda^2}$$

$$\sigma^2(X) = \overbrace{\mathbb{E}[X^2] - \mathbb{E}[X]^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$\sigma(X) = \frac{1}{\lambda}$$

Διάφορες εκθετικές πυκνότητες



Διασπορά γεωμετρικής κατανομής

$X \sim \Gamma\varepsilon\omega\mu(p)$, δηλ. X παίρνει τιμές $1, 2, \ldots$, και $\mathbb{P}[X = k] = p(1-p)^{k-1}$.

Έχουμε ήδη υπολογίσει $\mathbb{E}[X] = 1/p$.

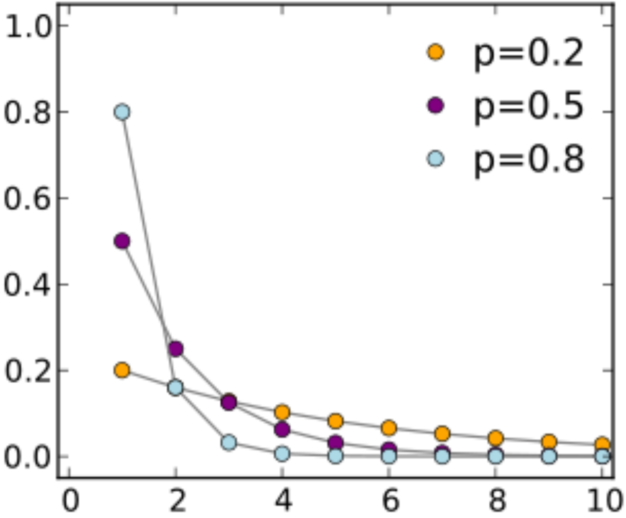
$$\sigma^2(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2} \qquad \sigma(X) = \frac{\sqrt{1-p}}{p}$$

Για το $\mathbb{E}[X^2] = \sum_{k=1}^{\infty} k^2 p (1-p)^{k-1}$ ξεκινάμε με $\boxed{\frac{1}{1-q} = \sum_{k=0}^{\infty} q^k}$ (με $q = 1-p$).

$$\frac{1}{(1-q)^2} = \sum_{k=0}^{\infty} k q^{k-1} \Rightarrow \frac{q}{(1-q)^2} = \sum_{k=0}^{\infty} k q^k \Rightarrow \frac{(1-q)^2 + 2(1-q)q}{(1-q)^4} = \sum_{0}^{\infty} k^2 q^{k-1} \Rightarrow$$

$$\Rightarrow \frac{1-q+2q}{(1-q)^3} = \sum k^2 q^{k-1} \Rightarrow \frac{2-p}{p^3} = \sum k^2 (1-p)^{k-1} \Rightarrow \frac{2-p}{p^2} = \sum_{0}^{\infty} k^2 p (1-p)^{k-1}$$

Διάφορες γεωμετρικές κατανομές



Η κατανομή Poisson(λ)

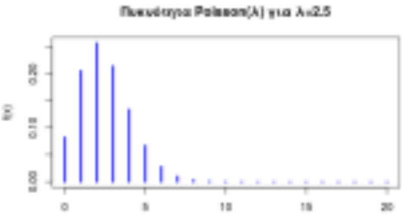
$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad (x \in \mathbb{R})$$

Αν $X \sim \text{Poisson}(\lambda)$ (με $\lambda > 0$) τότε

$$f_X(k) = e^{-\lambda} \boxed{\frac{\lambda^k}{k!}}, \qquad (k = 0, 1, 2, \dots)$$

Διακριτή κατανομή που συνήθως παριστάνει το πλήθος των τυχαίων γεγονότων που έχουν συμβεί σε κάποια χρονική περίοδο.

Π.χ., έχει μετρηθεί ότι το πλήθος από γκολ σε ένα αγώνα του Παγκοσμίου Κυπέλλου ποδοσφαίρου $\sim \text{Poisson}(2.5)$.



Η κατανομή Poisson(λ): μέση τιμή και διασπορά

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(\underbrace{k-1}_{m})!} = \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{\infty} k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$$

$m=k-2$

$$= \lambda^2 e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2$$

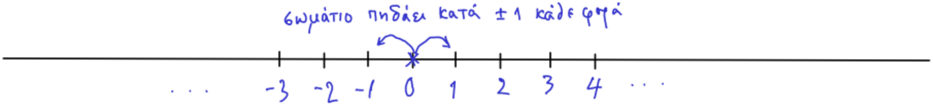
$$\mathbb{E}\left[X^2\right] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = \lambda^2 + \lambda$$

$$\sigma^2(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

$$\sigma(X) = \sqrt{\lambda}$$

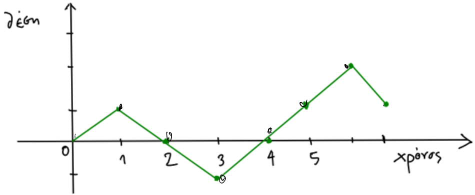
Ο συμμετρικός τυχαίος περίπατος στους ακεραίους

Σωμάτιο βρίσκεται στη θέση 0 στο χρόνο 1. Κάθε 1 sec πηδάει δεξιά ή αριστερά με ίση πιθανότητα.



Πηδήματα $X_1 = \pm 1, X_2 = \pm 1, \ldots$, ανεξάρτητες, ισόνομες ΤΜ, με

$$\mathbb{P}\left[X_j = 1\right] = \mathbb{P}\left[X_j = -1\right] = 1/2.$$



Θέση στο χρόνο n (με $S_0 = 0$):

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$\mathbb{E}\left[S_n\right] = \mathbb{E}\left[X_1\right] + \cdots \mathbb{E}\left[X_n\right] = 0.$$

Ο τυχαίος περίπατος: τυπική απόκλιση

$S_n = X_1 + \cdots + X_n$ ανεξάρτητα

$\sigma^2(X_1) = \mathbb{E} X_1^2 - \cancel{(\mathbb{E} X_1)^2}$

$\sigma^2(S_n) = \sigma^2(X_1) + \cdots + \sigma^2(X_n) = n\sigma^2(X_1)$ 1

Αφού $\mathbb{E}[X_1] = 0$ έχουμε $\sigma^2(X_1) = \mathbb{E}[X_1^2] = \mathbb{E}[1] = 1$, άρα

$$\sigma(S_n) = \sqrt{n}.$$

Πληρέστερη πληροφορία: $\frac{S_n}{2} + \frac{n}{2} \sim B(n, 1/2)$ (διωνυμική κατανομή).

$$\frac{S_n}{2} \sim -\frac{n}{2} + B(n, 1/2)$$

διορθωμένο από το video