Kom ihåg: $n$ ober. exp. lyckas/missl. m.s. $p/1-p$ och $X = \#$ lyckade försök $\Rightarrow X \sim Bin(n,p)$

s.a. $$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k=0,1,\ldots,n.$$

---

Försäkringsbolag försäkrar $n$ hus mot brand.

$X = \#$ nedbrunna hus $\quad X \sim Bin(n,p)$ där $p$ = sann. för brand.

Tal: David jobbar på Folksam och har hand om 1000 försäkrade hus. Om $p = 0.0001$, vad är sann att högst 2 hus brinner upp?

L: $$\mathbb{P}(X \leq 2) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2)$$

$$= \binom{n}{0} p^0 (1-p)^n + \binom{n}{1} p (1-p)^{n-1} + \binom{n}{2} p^2 (1-p)^{n-2}$$

$$= 0.9999^{1000} + n \cdot 0.0001 \cdot 0.9999^{999} + \frac{n(n-1)}{2} 0.0001^2 \cdot 0.9999^{998}$$

$$\cancel{\approx 0.9998} \approx 0.99984578$$

---

Funkar men vad händer om David har $n = 1000000$ hus och vi undrar

$$\mathbb{P}(X \leq 35) = \binom{n}{0} p^0 (1-p)^n + \cdots + \binom{n}{35} p^{35} (1-p)^{n-35}$$

jobbigt!

Då använder vi Poisson-fördelningen!

Uppkommer i liknande situationer som $\boxed{F5}$ 2

Binomial med <u>n stort och p litet.</u>

<u>Ex1:</u> # olyckor i en vägkorsning per månad (p. 213)

(n = # bilar i korsn. p = sann en bil har en olycka)

<u>Ex2:</u> # defekter hos nya bilar

(n = # nya bilar, p = sann. för en defekt)

<u>Ex3:</u> # stulna cyklar som Folksam måste betala

för under ett år. (n = ? p = ?)

<u>Ex4:</u> # ankommande till akuten en fredagskväll.

(n = ? p = ?)

<u>Ex5:</u> # som ringer datasupport under en vecka.

(n = ? p = ?)

---

<u>OBS:</u> Kan vara svårt att avgöra när Poisson-

förd. är lämplig.

---

Vad utmärker en Poisson-fördelning?

1) Experimentet består av att räkna # händelser

under en viss tid / i en viss population / i en viss volym

2) Sann för en händelse inträffar i ett visst

tidsinterval / del av pop. / i en volymsenhet är <u>samma.</u>

3) # händelser som inträffar i ett visst tidsintervall (del av pop. / volym ...) är oberoende av # händelser i alla andra tidsintervall.

---

För att kunna räkna behöver vi slt.

Def: $X$ är Poisson-fördelad med parameter $\lambda > 0$ om

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{för } k = 0, 1, 2, \ldots$$

Anm: Vi skriver $X \sim \mathrm{Poi}(\lambda)$ och vi har att $\mathbb{E}[X] = \lambda$ och $\mathrm{Var}(X) = \lambda$.

---

Anm: Om vi använder Poisson-förd. för att approximera Binomial har vi att $\lambda = np$.

---

Tal (David): Låt nu $X \sim \mathrm{Poi}(\lambda)$ där $\lambda = np = 1000 \cdot 0.0001 = 0.1$. Vi har att

$$\mathbb{P}(X \leq 2) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2)$$

$$= e^{-0.1} + 0.1 \cdot e^{-0.1} + \frac{0.1^2}{2} \cdot e^{-0.1} \approx 0.99984535 //$$

---

Skillnaden mellan detta och vårt förra svar är $\approx 0.0000005$. Jättelitet.

# Något mer om väntevärden

Kom ihåg att $\mathbb{E}[X] = \sum x\, \mathbb{P}(X = x)$

summa över alla möjliga värden för $X$.

Vi har följande:

1) Om $X, Y$ är 2 s.v. gäller att

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

2) Om $X_1, X_2, X_3, \ldots, X_n$ har samma fördelning och $\mu = \mathbb{E}[X_1] = \mathbb{E}[X_2] = \cdots = \mathbb{E}[X_n]$

då gäller att $\mathbb{E}\left[\frac{X_1 + \cdots + X_n}{n}\right]$

$$= \frac{1}{n}\mathbb{E}[X_1 + \cdots + X_n] \overset{1)}{=} \frac{1}{n}\left(\mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]\right)$$

$$= \frac{1}{n}(\mu + \cdots + \mu) = \frac{n\mu}{n} = \mu$$ (viktigt i statistik)

---

Ex:
Johanna tränar frisparkar på träning. Hon skjuter 100 st och sann att en viss frispark går i mål är 0.1 (och samma för alla 100 frisparkar). Vad är det förväntade # mål?

L: Låt

$$X_1 = \begin{cases} 1 & \text{om frispark 1 leder till mål} \\ 0 & \text{annars.} \end{cases}$$

I allmänhet

$$\mathbb{X}_k = \begin{cases} 1 & \text{om frispark } k \text{ leder till mål.} \\ 0 & \text{annars.} \end{cases}$$

Om $\mathbb{X} = \mathbb{X}_1 + \mathbb{X}_2 + \dots + \mathbb{X}_{100}$ vad är $\mathbb{X}$?

Jo $\mathbb{X} = \#$ mål totalt.

Vi söker $\mathbb{E}[\mathbb{X}] = \mathbb{E}[\mathbb{X}_1 + \dots + \mathbb{X}_{100}] = \mathbb{E}[\mathbb{X}_1] + \dots + \mathbb{E}[\mathbb{X}_{100}]$.

Vidare är $\mathbb{E}[\mathbb{X}_k] = 0 \cdot \mathbb{P}(\mathbb{X}_k = 0) + 1 \cdot \mathbb{P}(\mathbb{X}_k = 1)$

$= 0 + p = p = 0.1$

$\Rightarrow \quad \mathbb{E}[\mathbb{X}] = p + p + \dots + p = 100 \cdot p = 100 \cdot 0.1 = 10$ //

---

Mer intressant (?):

1) om $\mathbb{X} = \mathbb{X}_1 + \dots + \mathbb{X}_n$ har vi att $\mathbb{E}[\mathbb{X}] = np$

2) Vilken fördelning har $\mathbb{X}$?

Vi har $n$ försök. $\mathbb{X}_k = 1$ betyder lyckat

$\mathbb{X}_k = 0$ " misslyckat.

$\Rightarrow \quad \mathbb{X} = \mathbb{X}_1 + \dots + \mathbb{X}_n = \#$ lyckade försök.

$\Rightarrow \quad \mathbb{X} \sim Bin(n, p)$.

D.v.s. Om $\mathbb{X} \sim Bin(n,p)$ gäller att $\mathbb{E}[\mathbb{X}] = np$.