Www-palvelimen mitoitus (10 min)

- Kuinka monta kyselyä per s pitää voida käsitellä, jotta tn(palvelin pois käytöstä) $\leq 1\% = p$?
- Kiirejakson aikana saapuu λ kyselyä per s.
  (Esim. 10^6 per päivä $\rightsquigarrow$ 10^5 per kiiretunti $\rightsquigarrow$ $\approx \frac{10^5}{3600} \approx 28$ per s)
- Olkoon d.s.m. $N =$ "lkm: kyselyt per s kiirejakson aikana"
  $\mathbb{E}N = \lambda$
- Halutaan etsiä luku k_0 s.e. $\mathbb{P}(N > k_0) \leq p = \frac{1}{100}$.
  Tällöin k_0 kyselyä per s kapasiteetin palvelin OK.
- Ongelmia:
  - Miten mallinnetaan N?
  - Miten lasketaan $\mathbb{P}(N > k_0)$?

2

Pienten lukujen laki (15 min)

- $N = \text{lkm}$(kyselyt www-palvelun valitun 1s aikajakson kuluessa)
- $A_i := \{$"asiakas i tuottaa kyselyn"$\}$, $i = 1, \dots, n$.
  (esim. VR: $n \approx 3 \cdot 10^6$)
- $N = \sum_{i=1}^{n} \Theta_i$
- (järkevä) oletus: $\Theta_1, \dots, \Theta_n$ riippum.
- $\Theta_i \sim \text{Ber}(p)$, mikä on p?
- Oletus: $\lambda = \mathbb{E}N = np \Rightarrow p = \lambda/n$.

$\Rightarrow N \sim \text{Bin}(n, \lambda/n)$

$$P(N=k) = \binom{n}{k} (\lambda/n)^k (1-\lambda/n)^{n-k}, \quad k = 0, \dots, n.$$

Koska n hyvin iso, arvioidaan: $n \approx \infty$.

$$\lim_{n\to\infty} P(N=k) = ? \qquad (k \in \mathbb{Z}_+).$$

- $$P(N=k) = \frac{\lambda^k}{k!} \left[\frac{n!}{(n-k)!\, n^k}\right] \frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}$$
- $$[\dots] = \frac{n(n-1)\cdots(n-k+1)}{n^k} = \prod_{j=0}^{k-1} \left(\frac{n-j}{n}\right) = \prod_{j=0}^{k-1} (1 - j/n) \longrightarrow 1$$
- $(1-\lambda/n)^k \longrightarrow 1$. $x = -\lambda$
- $(1-\lambda/n)^n = (1 + \frac{x}{n})^n \longrightarrow e^x = e^{-\lambda}$

Siis raja: $P(N=k) \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}$

<u>Määr</u> Poisson-jak. param. $\lambda > 0$ on tnf $p_\lambda(k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k \geq 0$.
(D.s.m. X on $\text{Poi}(\lambda)$ jos sen ptnf on $\text{Poi}(\lambda)$.)

<u>Huom.</u> $\sum_{k \geq 0} p_\lambda(k) = 1$, joten p_λ on tnf $\mathbb{Z}$:ssa.

<u>Lause (PLL)</u> Jos $N_n \sim \text{Bin}(n, \lambda/n)$ kaikilla $n \geq 1$, niin

$$\lim_{n\to\infty} P(N_n = k) = P(X=k) \quad \forall k \geq 0,$$

missä $X \sim \text{Poi}(\lambda)$.

Poisson-jakauman o.a. (5 min)

$X \sim \mathrm{Poi}(\lambda)$.

$\mathbb{E}X = ?$

$$\mathbb{E}X = \sum_{k=0}^{\infty} k\,\mathbb{P}(X=k) = \sum_{k=0}^{\infty} k\, e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

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

$$= \lambda \sum_{k=0}^{\infty} \mathbb{P}(X=k)$$

$$= \lambda\, \mathbb{P}(X \in \{0,1,\dots\})$$

$$= \lambda$$

(3

Mutta miten lasketaan

$\mathbb{P}(X > k)$, kun $X \sim \mathrm{Poi}(28)$?

D.s.m:n varianssi (15 min)

__Määr__ $X : \Omega \to S \subset \mathbb{R}$ d.s.m. tn-avr:llä $(\Omega, 2^\Omega, P)$. Jos $\mu = \mathbb{E}X$ on olemassa, niin X:n varianssi on $\sigma_X^2 = \mathrm{Var}(X) := \mathbb{E}(X-\mu_X)^2$.

__Laskeminen__:

$$\mu_X = \mathbb{E}X = \sum_{k \in S} k\, p_X(k).$$

$$\sigma_X^2 = \underset{f(s)=(s-\mu_X)^2}{\mathbb{E}f(X)} = \sum_{k\in S} f(k)p_X(k) = \sum_{k\in S}(k-\mu_X)^2 p_X(k)$$

__Lause__ $\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.

__Tod__

$$\begin{aligned}\mathrm{Var}(X) &= \sum_{k\in S}(k^2 - 2\mu_X k + \mu_X^2)p_X(k) \\ &= \sum_{k\in S} k^2 p_X(k) - 2\mu_X \sum_{k\in S} k\,p_X(k) + \mu_X^2 \sum_{k\in S} p_X(k) \\ &= \mathbb{E}X^2 - 2(\mathbb{E}X)^2 + (\mathbb{E}X)^2. \quad \square\end{aligned}$$

Varianssin summakaava (20 min)

Lause Olkoot X, Y dsm jolla $\mu_X, \mu_Y \in \mathbb{R}$ ja $a \in \mathbb{R}$.

(i) $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$

(ii) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$, jos $X \perp\!\!\!\perp Y$.

(iii) $\mathrm{Var}(a) = 0$

Tod. (i) $\mathrm{Var}(aX) = \mathbb{E}(aX)^2 - (\mathbb{E}(aX))^2$

$= a^2 \mathbb{E}(X^2) - a^2(\mathbb{E}X)^2 = a^2 \mathrm{Var}(X)$.

(ii) $\mathbb{E}((X+Y)^2) = \mathbb{E}X^2 + 2\mathbb{E}XY + \mathbb{E}Y^2$

$(\mathbb{E}(X+Y))^2 = (\mathbb{E}X)^2 + 2\mathbb{E}X\mathbb{E}Y + (\mathbb{E}Y)^2$

$\quad \| $

$(\mathbb{E}X + \mathbb{E}Y)^2 =$

$\Rightarrow \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2(\mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y)$

$= \mathrm{Var}(X) + \mathrm{Var}(Y)$

(iii) $\mathbb{E}a^2 - (\mathbb{E}a)^2 = a^2 - a^2 = 0$.

Esim. • $\Theta \sim \mathrm{Ber}(p)$

$\mathrm{Var}(\Theta) = \mathbb{E}\Theta^2 - (\mathbb{E}\Theta)^2$

$= p \cdot 1^2 + (1-p)0^2 - (p \cdot 1 + (1-p) \cdot 0)^2$

$= p - p^2$

$= p(1-p)$

• $B \sim \mathrm{Bin}(n,p)$: $B = \Theta_1 + \dots + \Theta_n$: $\perp\!\!\!\perp \sim \mathrm{Ber}(p)$

$\mathrm{Var}(B) = \mathrm{Var}(\Theta_1) + \dots + \mathrm{Var}(\Theta_n) = np(1-p)$

• $X \sim \mathrm{Poi}(\lambda)$: $\mathbb{E}X = \lambda$

$\mathbb{E}X^2 = \sum_0^\infty k^2 e^{-\lambda} \frac{\lambda^k}{k!} = \sum_1^\infty \dots = \sum_1^\infty k(k-1) e^{-\lambda} \frac{\lambda^k}{k!} + \sum_{k=1}^\infty k e^{-\lambda} \frac{\lambda^k}{k!}$

$= \sum_2^\infty e^{-\lambda} \frac{\lambda^k}{(k-2)!} + \mathbb{E}X$

$= \lambda^2 + \lambda \Rightarrow \mathrm{Var}(X) = (\lambda^2 + \lambda) - \lambda^2 = \lambda$.

Harvinaisten tapahtumien tn-estimointi (15 min)

Lause (Markovin ey)

Olkoon $X : \Omega \to S \subset \mathbb{R}$ d.s.m. s.e. $X(\omega) \geq 0 \;\forall \omega \in \Omega$. Tällöin

$$P(X \geq a) \leq \frac{1}{a} \mathbb{E}X \quad \forall\, a > 0.$$

Tod.

$$\mathbb{E}X = \sum_{x \in S} x\, P(X = x) \geq \sum_{x \in S:\, x \geq a} x\, P(X = x)$$

$$\geq a \sum_{x \in S:\, x \geq a} P(X = x) = a\, P(X \geq a).$$

Esim. $X \sim \mathrm{Poi}(28)$, $\mathbb{E}X = 28$

$$P(X \geq 100) \leq \frac{1}{100} \mathbb{E}X = 28\,\%.$$

Tarkennus:

Lause (Chebyshevin ey)

Olkoon $X : \Omega \to S \subset \mathbb{R}$ d.s.m. s.e. $\mu_X = \mathbb{E}X \in \mathbb{R}$. Tällöin

$$P(|X - \mu_X| \geq a) \leq \frac{1}{a^2} \mathrm{Var}(X) \quad \forall\, a > 0.$$

Tod. $Y(\omega) := (X(\omega) - \mu_X)^2$.

Tällöin $P(|X - \mu_X| \geq a) = P(Y \geq a^2)$

$$\underset{\text{Markov}}{\leq} \frac{1}{a^2} \mathbb{E}Y = \frac{1}{a^2} \mathrm{Var}(X)$$

Esim. $X \sim \mathrm{Poi}(28)$; $\mathbb{E}X = 28$; $\mathrm{Var}(X) = 28$

$$P(X \geq 100) = P(X - \mu_X \geq 72)$$

$$\leq P(|X - \mu_X| \geq 72)$$

$$\leq \frac{1}{72^2} \mathrm{Var}(X) = \frac{28}{72^2}$$

... Www-palvelimen mitoitus

- Mallinnetaan $N \sim Poi(\lambda)$; $\lambda = 28$
- Etsitään k_0 s.e. $P(N > k_0) \leq p := \frac{1}{100}$.
- $EN = \lambda$
- $Var(N) = \lambda$

Kun $k > \lambda$, niin

$$P(N>k) = P(N-\lambda > k-\lambda)$$
$$\leq P(|N-\lambda| > k-\lambda)$$
$$\stackrel{\text{Cheb.}}{\leq} \frac{1}{(k-\lambda)^2} Var(N)$$
$$= \frac{\lambda}{(k-\lambda)^2}$$

$$\frac{\lambda}{(k-\lambda)^2} = p:$$

$$53^2 = 2809$$

$$k^2 - 2\lambda k + \lambda^2 = \lambda/p$$
$$k^2 - 2\lambda k + (\lambda^2 - \lambda/p) = 0$$
$$k = \frac{2\lambda \pm \sqrt{4\lambda^2 - 4(\lambda^2 - \lambda/p)}}{2} = \frac{2\lambda \pm \sqrt{4\lambda/p}}{2}$$
$$= \lambda \pm \sqrt{\lambda/p}$$
$$= 28 \pm \sqrt{2800}$$

$$k > \lambda \Rightarrow k = \lambda + \sqrt{\lambda/p} \approx \underbrace{81}$$

$\lambda = 28$, $p = \frac{1}{100}$

Siis: kun $k \geq k_0 = 81$, niin $\frac{\lambda}{(k-\lambda)^2} \leq \frac{1}{100}$.

$\Rightarrow P(N > 81) \leq 1\%$

$\Rightarrow$ 81 palvelukapasiteettiä riittää.