Bernoulli RV

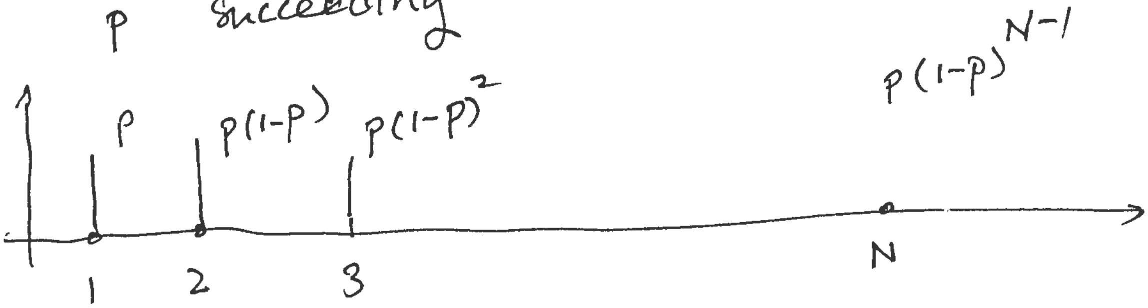
1	p
0	(1-p)

$E[X] = p$

$E[(X - E[X])^2] = p(1-p)$

Geometric distribution

p succeeding



$E[X] = \frac{1}{p}$

$VAR[X] = \frac{1-p}{p}$

Binomial

$P(H) = p$

flip N times

H heads, $N-H$ tails

$$\frac{N!}{H!\,(N-H)!}\, p^{H} (1-p)^{N-H} \longrightarrow P_b(H;\, N, p)$$

Mean: $E[X] = Np$

$\mathrm{var}[X] = Np(1-p)$

Multinomial

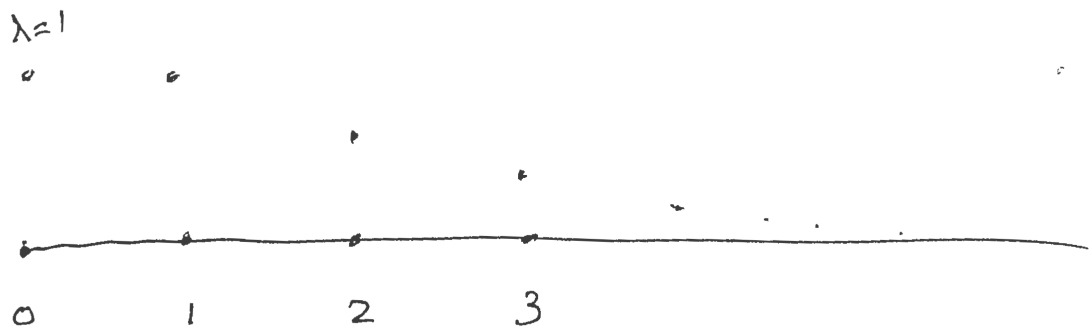
$$P(n_1, n_2 \cdots n_K; p_1, p_2 \cdots)$$

$$= \frac{N!}{n_1! \cdots n_K!} p_1^{n_1} \cdots p_K^{n_K}.$$

$$\sum_{i=1}^{\infty} i\,p(1-p)^{(i-1)} = p \sum_{i=1}^{\infty} i(1-p)^{i-1}$$

$$= p\Big[\sum_{i=1}^{\infty} (1-p)^{i-1} + \sum_{i=2}^{\infty} (1-p)^{i-1} + \cdots$$

Poisson distribution



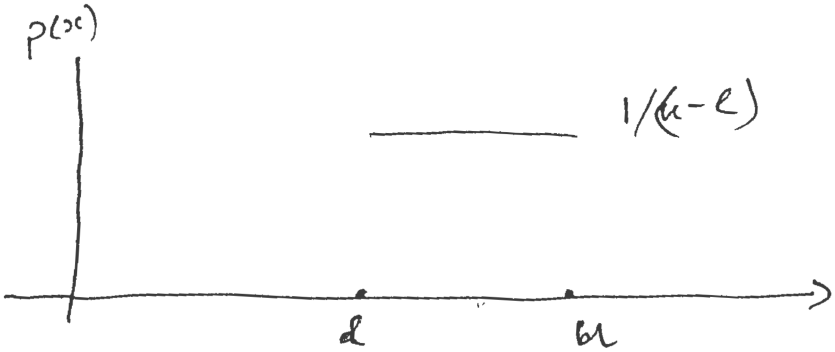
$$P(\{X=k\}) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\sum_{k=0}^{\infty} P(\{X=k\}) = 1 = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \boxed{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}$$

$$E[X] = \lambda$$

$$E[(X - E[X])^2] = \lambda$$

Uniform



Beta - distribution

$$P_\beta(x \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{(\alpha-1)} (1-x)^{(\beta-1)}$$

$0 \leq x \leq 1$

$P_\beta(x|\alpha,\beta)$ for $\alpha > 1, \beta > 1$ has a single maximum at $x = (\alpha-1)/(\alpha+\beta-2)$.

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

$$E[(X - E[X])^2] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Gamma dist

$$P_\gamma(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot x^{(\alpha-1)} e^{-\beta x}$$

$x \geqslant 0$

$\alpha > 0$

$\beta > 0$

mean is $\frac{\alpha}{\beta}$

variance is $\frac{\alpha}{\beta^2}$

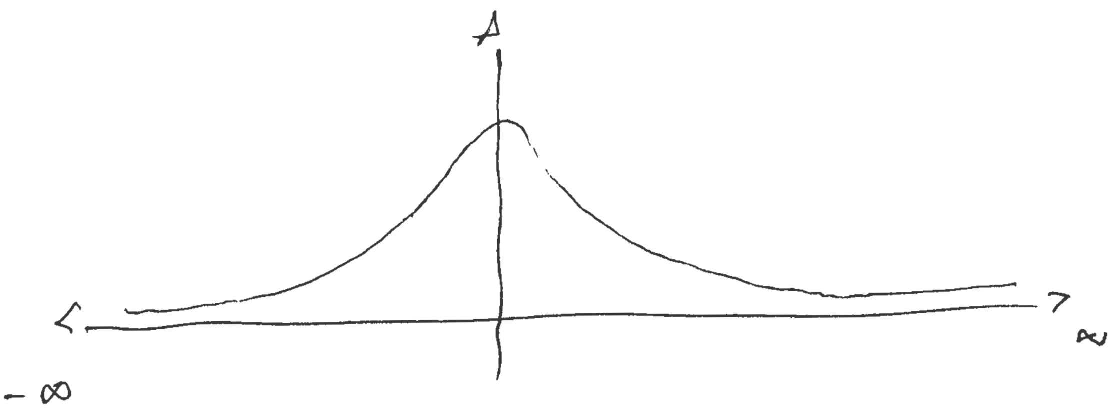
Exponential dist

$x \geq 0 \qquad \lambda e^{-\lambda x}$

otherwise $\qquad 0$

Standard normal dist

$$p(x) = \left(\frac{1}{\sqrt{2\pi}}\right) e^{-\frac{x^2}{2}}$$



$$\int_{-\infty}^{\infty} x\, p(x)\, dx = 0 = \text{mean}$$

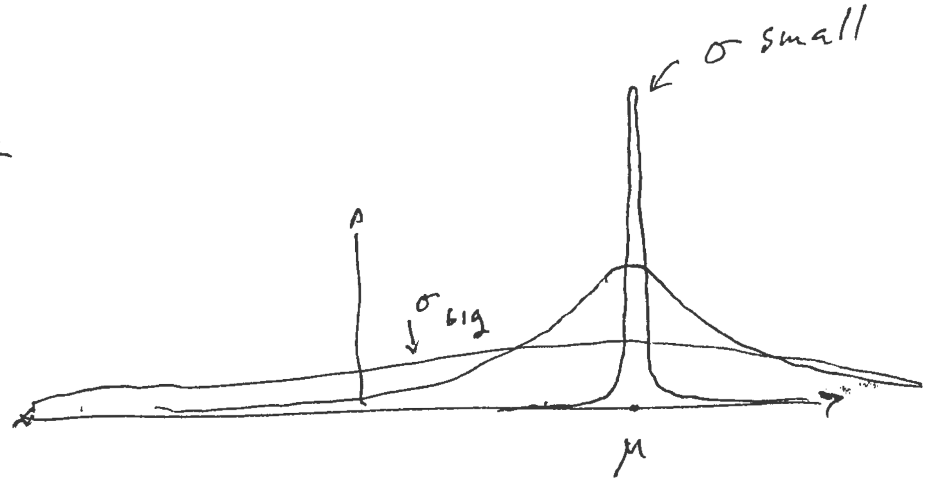
$$\text{Var}(x) = 1$$

Normal dist

$$p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

mean μ

variance σ^2



$$\int_{\mu-\sigma}^{\mu+\sigma} p(x)\, dx = P(\{X \in [\mu-\sigma, \mu+\sigma]\})$$

$\approx .68$

$\mu+2\sigma$ $\approx .95$

$\mu-2\sigma$

3σ $\approx .99$

3σ

$$x = \frac{h - Np}{\sqrt{Np\underbrace{(1-p)}_{q}}}$$

$$P(\{x \in [a,b]\}) \approx \int_a^b \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

$p = 0.7$ $\quad$ $(1-p) = 0.3$ $\quad$ $\sqrt{p(1-p)} \approx 0.5$

$N = 1e^6$ $\quad$ $h \in \{1e3, 2e3\}$

$$\int_{\frac{1e3 - 1e6 \cdot 0.7}{\sqrt{1e6 \times 0.3 \times 0.7}}}^{\frac{2e3 - 1e6 \times 0.7}{\sqrt{1e6 \times 0.3 \times 0.7}}} \frac{1}{\sqrt{2\pi}} \cdot e^{-u^2/2}\, du$$

$N = 1e6 \qquad \{ h < 1e^5 \}$

$$x = \frac{h - Np}{\sqrt{Np(1-p)}}$$

$$h < 1e^5 \quad \Rightarrow \quad x < \frac{1e^5 - 1e^6 \cdot 0.7}{\sqrt{1e6 \times 0.7 \times 0.3}}$$

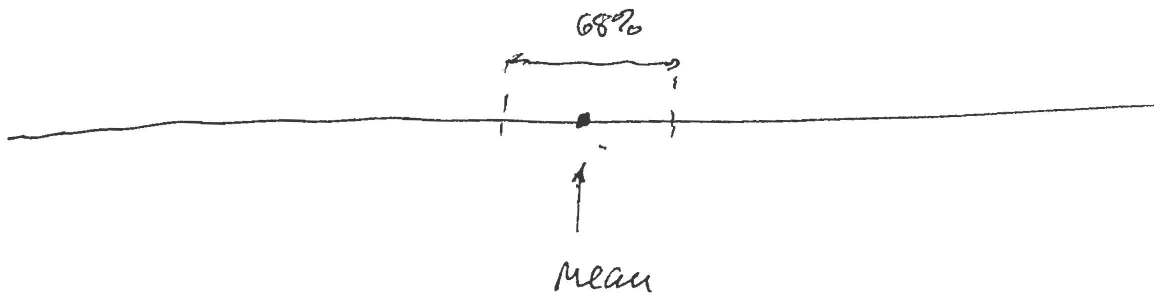
$$\int_{-\infty}^{(\quad)} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du$$

N big

p

~~h = Np~~

$$h \approx \text{Normal}\Big(\underset{\substack{\uparrow\\ \text{mean}}}{Np}\ ,\ \sqrt{Np(1-p)}\Big).$$



$$\frac{h}{N} \approx \text{Normal}\Big(\underset{\substack{\uparrow\\ \text{mean}}}{p}\ ,\ \underset{\substack{\uparrow\\ \text{std}}}{\sqrt{\frac{p(1-p)}{N}}}\Big)$$


k
k+1

$$P(X_{k+1} = j \mid X_k = i)$$

$$= \underline{P_{ij}}$$

$$X_0 = 1$$

$$P(X_1 = j \mid X_0 = 1) = P_{1j}$$

$$P(X_0 = i) = \underline{\pi_i}$$

$$P(X_1 = j) = \sum_i P(X_1 = j \mid X_0 = i) P(X_0 = i)$$

$$= \sum_i \pi_i P_{ij} = \underline{\pi} \, \underline{\underline{P}}$$

$$P(X_2 = j) = \sum_k P(X_2 = j \mid X_1 = k) \cdot P(X_1 = k)$$

$$= \sum_k P_{kj} \left[\sum_i \pi_i P_{ik} \right]$$

$$= \left[\underline{\pi} \, \underline{\underline{P}} \right] \underline{\underline{P}}$$

$$\left[\Pi\ \underline{\underline{P}}^{N}\right]$$

$$\uparrow$$

$$\underline{p}_s = \underline{p}_s\,\underline{\underline{P}}$$

X is uniform $\in [0-1]$

Y is uniform $[0-10]$

indep.

$$Z = (Y-5X)^3 - X^2$$

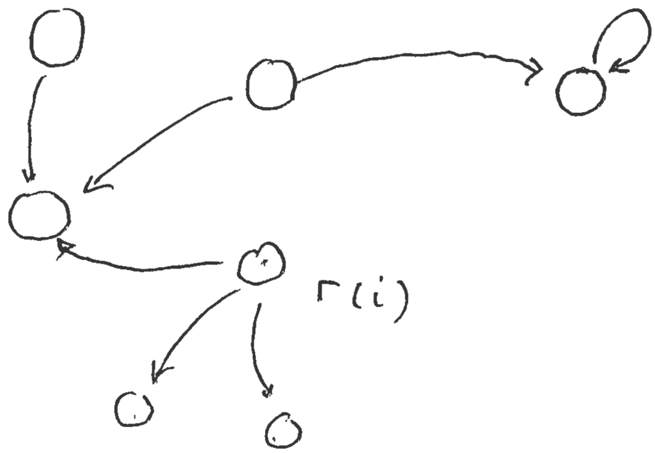
$$Var(Z) = E[Z^2] - E[Z]^2$$

$$2.76 \times 10^4$$

$$P(\{Z > 3\}) \approx 0.62$$

$$P(\{Z > 950\})$$

$$\underline{\text{Std}} \approx \frac{C}{\sqrt{N}}$$



$$r(j) = \sum_{i \to j} \frac{r(i)}{|i|} = \sum P_{ij} \quad r(i)$$

$$\frac{1}{n} \cdot \underline{1}$$

$$(1-\alpha)\frac{1}{|i|} \quad + \quad \alpha\left(\underline{1}\,\underline{1}^T\right)$$

The maximum likelihood principle

D - data

$p(D|\theta)$ - model

θ - models parameters

choose $\hat{\theta}$ such that $p(D|\hat{\theta})$ is a maximum

$$p(D|\theta) = \mathcal{L}(\theta)$$

N times

K heads

$\Theta = p(\text{head})$

$$p(D|\Theta) = \left(\frac{N!}{K!\,(N-K)!}\right)\Theta^{(K)}(1-\Theta)^{(N-K)}$$

$$\left(\frac{N!}{K!\,(N-K)!}\right)\left[K\,\Theta^{(K-1)}\cdot(1-\Theta)^{(N-K)} - (N-K)\,\Theta^{(K)}(1-\Theta)^{(N-K-1)}\right]$$

$$K(1-\Theta) - (N-K)\,\Theta = 0$$

$$\hat{\Theta} = \cancel{\frac{N}{K}}\ \frac{K}{N}$$

N times, stopping at first head

$\theta = p(H)$.

$$p(D|\theta) = (1-\theta)^{(N-1)}\,\theta$$

$$(1-\theta)^{(N-1)} - (N-1)(1-\theta)^{(N-2)} \cdot \theta$$

$$\hat{\theta} = \frac{1}{N}$$

Data $= N_1, N_2$

$(N_1 - 1)T$, followed by 1 head

$p(H)$

$$P(N_1, N_2 \mid \theta) = p(N_1 \mid \theta)\, p(N_2 \mid \theta)$$

$$= \left[(1-\theta)^{N_1 - 1} \cdot \theta\right]\left[(1-\theta)^{(N_2 - 1)} \cdot \theta\right]$$

$$= \left[(1-\theta)^{(N_1 + N_2 - 2)} \cdot \theta^2\right]$$

$$-(N_1 + N_2 - 2)(1-\theta)^{(N_1+N_2-3)}\theta^2 + 2\theta(1-\theta)^{(N_1+N_2-2)} = 0$$

$$\cancel{(N_1+N_2-2)}$$

$$\theta(N_1 + N_2 - 2) = 2(1-\theta)$$

$$\hat{\theta} = \left[\frac{2}{N_1 + N_2}\right]$$

$$\underline{\text{Data}} \quad \overbrace{(N_1, h_1)}^{D_1}, \quad \overbrace{(N_2, h_2)}^{D_2}$$

$$p(D_1, D_2 \mid \theta) = p(D_1 \mid \theta)\, p(D_2 \mid \theta)$$

$$= \left[(\text{stuff})\, \theta^{h_1} (1-\theta)^{N_1 - h_1} \right] \left[(\text{stuff})\, \theta^{h_2} (1-\theta)^{N_2 - h_2} \right]$$

$$= \left[(\text{stuff})\, \theta^{(h_1+h_2)} \cdot (1-\theta)^{(N_1+N_2-(h_1+h_2))} \right]$$

$$\hat{\theta} = \frac{h_1 + h_2}{N_1 + N_2}$$

Data:

1hr $\quad n_1, n_2, n_3$

$$P(n_1 | \theta) = \frac{\theta^{n_1} \cdot e^{-\theta}}{n_1!}$$

$$P(n_1, n_2, n_3 | \theta) = P(n_1|\theta) \cdot P(n_2|\theta) \cdot P(n_3|\theta)$$

$$\log P(n_1, n_2, n_3 | \theta) = \log P(n_1|\theta) + \log P(n_2|\theta) + \log P(n_3|\theta)$$

$$\begin{aligned} &= (n_1 \log\theta - \theta) - \log(n_1!) \\ &\quad + (n_2 \log\theta - \theta) \\ &\quad + (n_3 \log\theta - \theta) \quad + \text{const} \\ &= (n_1+n_2+n_3)\log\theta - 3\theta \quad + \text{const} \end{aligned}$$

$$\hat{\theta} = \cancel{\frac{3}{n_1+n_2+n_3}} = \frac{n_1+n_2+n_3}{3}$$

Data:

$$x_1 \ldots, x_N$$

$$p(x|\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-(x-\theta)^2/2\sigma^2}$$

$$p(x_1, x_2 \ldots x_N|\theta) = p(x_1|\theta)\, p(x_2|\theta) \ldots p(x_N|\theta)$$

$$\log p(x_1, x_2 \ldots x_N|\theta) = \log p(x_1|\theta) + \log \ldots$$

$$= -\frac{(x_1-\theta)^2}{2\sigma^2} - \frac{(x_2-\theta)^2}{2\sigma^2} - \ldots + \text{'C}$$

care about : $-\sum_i (x_i-\theta)^2$

get $\hat{\theta} = \frac{\sum_i x_i}{N}$