$P(H) = p$ 1st trial

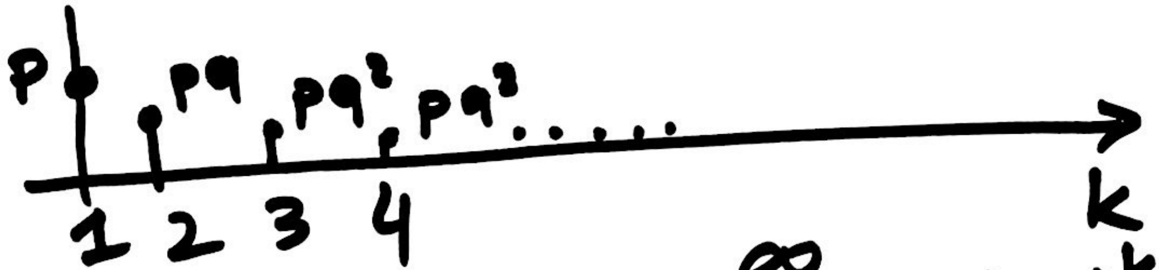
$P(TH) = p(1-p)$ 2nd " $q = 1-p$

$P(TTH) = (1-p)^2 p$ 3rd "

$\vdots$

$P(TT\ldots TH) = (1-p)^{k-1} p$ k-th trial

p.m.f. Geometric



$$E(X) = \sum_{k=1}^{\infty} k \, P_r(X=k) = \sum_{k=1}^{\infty} k \, p \, (1-p)^{k-1} = \frac{1}{p}$$

$$\left(\alpha^0 \quad 1 + \alpha + \alpha^2 + \alpha^3 + \ldots = \frac{1}{1-\alpha} \right)$$

$$E[X] = \sum_{k=1}^{\infty} k(1-q)\, q^{k-1} = \sum_{k=1}^{\infty} k q^{k-1} - k q^{k} \qquad (q = 1-p)$$

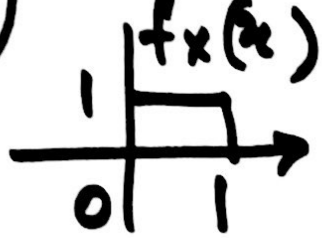
$$= \sum_{k=0}^{\infty} (k+1) q^{k} - (k+1) q^{k+1} = \sum_{k=0}^{\infty} q^{k} = \frac{1}{1-q}$$

$$= \frac{1}{p}$$

Scanned by CamScanner

$U[0,1]$

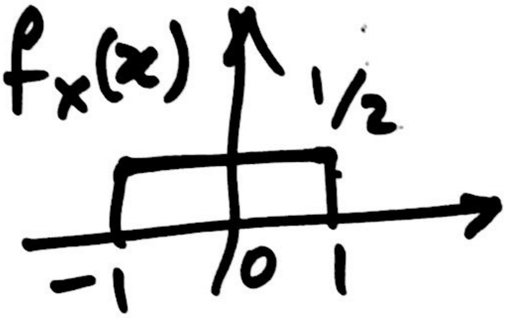
$$E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$



$$E[X] = \int_0^1 x \, 1 \, dx$$

$$E[X] = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$$

$U[-1,1]$



Average value

$E[X] = 0$

$$E[2X] = 2E[X]$$

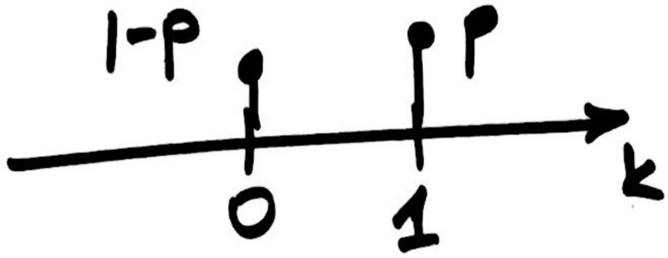
$$E[3X+5] = 3E[X] + 5$$

$$E[3X+5] = \int_{-\infty}^{\infty} (3x+5) f_X(x) \, dx$$

$$= 3\int x f_X(x) \, dx + \int 5 f_X(x) \, dx$$

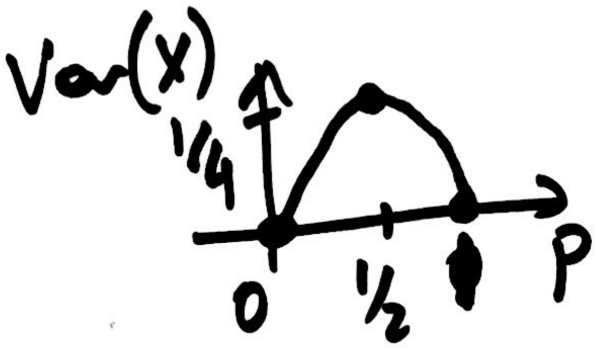
$$= 3E[X] + 5$$

p.m.f. Bernouilli



$$E[X^2] = 0^2(1-p) + 1^2 p = p$$

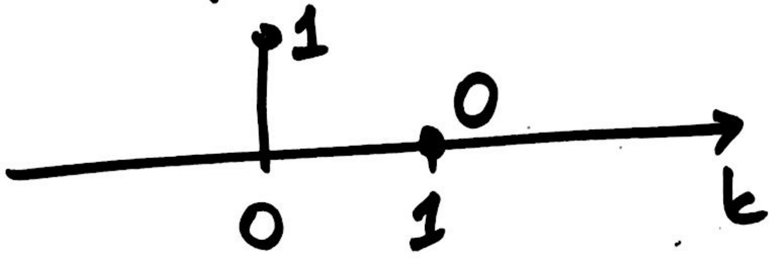
$$Var(X) = (0-\underbrace{p}_{=E(X)})^2(1-p) + (1-\underbrace{p}_{=E[X]})^2 \cdot p = p(1-p)$$



When $p=0$ Deterministic case

No variance

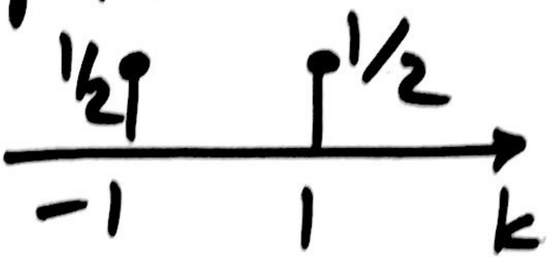
$Var(X)=0$



$$\sigma_X^2 = Var(X)$$

std. deviation $= \sigma$

pmf of X_i:



$E\{X_i\} = 0$

$Var(X_i) = (-1)^2 \frac{1}{2} + 1^2 \frac{1}{2} = 1$

$Var(aX+b) = a^2 Var(X)$

Multinomial = vector form of binomial

Poisson Example

\# of calls arriving a switchboard

5 calls per second

$\lambda = 5$ (Mean value)

Prob. of no calls in 1 second:

$$P(X=0) = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-\lambda} = e^{-5}$$

1 call per second)

$$P(X=1) = e^{-\lambda}\frac{\lambda^1}{1!} = 5e^{-5}$$

⋮

Poisson with a time parameter.

$$P(Y=k) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}, \quad k=0,1,\dots$$

Prob of No calls in 0.5 sec = T

$\lambda = 5$ calls/sec

$$P(Y=0) = \frac{(2.5)^0}{0!} e^{-2.5} = 0.082$$

$$\underbrace{\mu = np}_{\text{Binomial}} \qquad \sigma^2 = np(1-p)$$

Normal

Binomial has mean np
$B(n,p)$ variance $np(1-p)$

Set them to the mean & variance of the Normal.

$$K \sim B(n,p) \approx N(\underset{\mu}{\overset{\shortparallel}{np}},\ \underset{\sigma^2}{\overset{\shortparallel}{np(1-p)}})$$

$$K \sim N(np,\ np(1-p))$$

$$\hat{p} = \frac{K}{n} \sim N\left(p,\ \frac{p(1-p)}{n}\right)$$

$$E[K] = np, \qquad E\left[\frac{K}{n}\right] = \frac{np}{n}$$

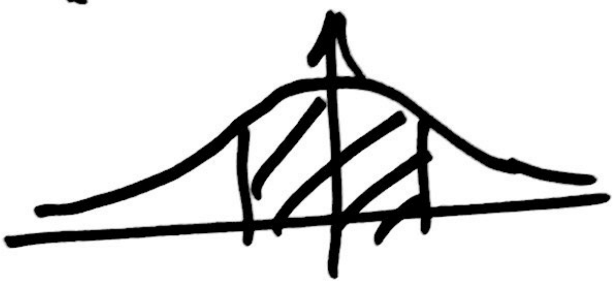
$$Var[K] = np(1-p) \qquad Var\left(\frac{K}{n}\right) = \frac{np(1-p)}{n^2}$$

Normal $X \sim N(\mu, \sigma^2)$

$P(X=3)=0$

$P(X \le \mu) = \frac{1}{2}$

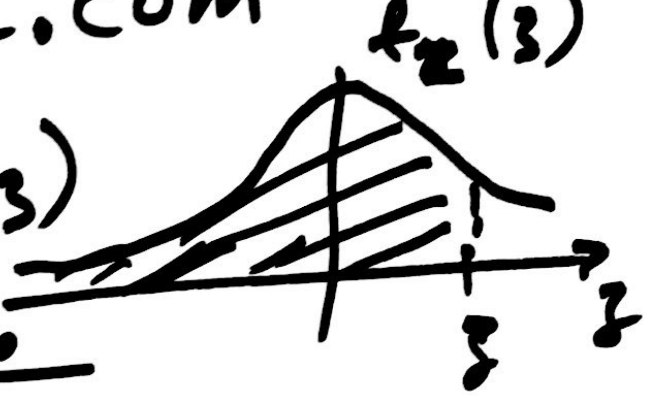
$P(\mu-\sigma < X < \mu+\sigma) \cong \frac{66}{100}$



z - tables. for $N(0, \overset{=\sigma^2}{1})$

onlinestatbook.com

$P(Z \le z) = F_Z(z)$



	Prob
$z=-\infty$	0
$\vdots$	
$z=0$	$1/2$
$\vdots$	
$z \doteq 3$	≥ 0.99
$z=\infty$	$=\underline{1}$

$$\Pr\left(\begin{array}{c}\text{data} \\ k \text{ out of } n\end{array}\right) = \binom{n}{k} p^{k} (1-p)^{n-k}$$

$$\max_{\text{wrt } p} \cancel{\binom{n}{k}} p^{k} (1-p)^{n-k}$$

No p

$$\text{likelihood}(p) = p^{k} (1-p)^{n-k}$$

max $\quad$ " $\quad$ = max log likeli-hood

$$LL(p) = k \ln p + (n-k) \ln(1-p)$$

$$0 = \frac{\partial LL}{\partial p}(p) = 0 \Rightarrow p = \frac{k}{n}$$

$$L(19\text{ heads} \mid p_1 = 1) = 0$$

$$P(19\text{ heads} \mid p_2 = 0.1) = \binom{20}{19} p^{19} (0.9)^1$$

$$= \binom{20}{19} 0.1^{19} (0.9)^1$$

$$L(19\text{ heads} \mid p_2 = 0.1) = (0.1)^{19} (0.9)^1$$

<u>Laplace Smoothing</u> $k_a + k_b + k_c = n$

Original probabilities k_a, k_b, k_c
n observat.

<u>step 1</u> $\cancel{p_a+1}, \cancel{p_b+1}, \cancel{p_c+1}$

$$p_a = \frac{k_a+1}{n+3}, \; p_b = \frac{k_b+1}{n+3}, \; \frac{k_c+1}{n+3} = p_c$$

$$3 = V = \{a, b, c\}$$

$$1 = \frac{k_a+1}{n+3} + \frac{k_b+1}{n+3} + \frac{k_c+1}{n+3}$$

Instead of

$$p_a' = k_a/n, \; p_b' = \frac{k_b}{n}, \; p_c' = \frac{k_c}{n}$$

Bernoulli Distribution

$$P(X=k)=\begin{cases} p, & k=1\\ 1-p, & k=0\end{cases}$$

or

$$P(X=k)=p^{k}(1-p)^{1-k}=\begin{cases} p, & k=1\\ 1-p, & k=0\end{cases}$$

For $j=1$

$$\pi_1 \prod_{i=1}^{\#\text{ of pixels}=784} P_{1i}^{x_i}(1-P_{1i})^{1-x_i} \qquad x_i=\begin{Bmatrix}1\\ \text{or}\\ 0\end{Bmatrix}$$

prob. of jth class

⋮

$j=10$

$$\pi_{10} \prod_{i=1}^{784} P_{10i}^{x_i}(1-P_{10i})^{1-x_i}$$

Observ. vector $x=(x_1, x_2, \dots, x_{784})$

image $\xrightarrow{\text{construct}}$ x

Model : Bernoulli

Class 1: spam

$$NB_1 = \pi_s \underbrace{\prod_{t=1}^{|V|=6} P_{1t}^{x_t} (1-P_{1t})^{1-x_t}}_{P(Doc|spam)}$$

$P(C=spam)=\pi_s$

Class 2: ham

$$NB_2 = \underbrace{\pi_h}_{P(ham)} \underbrace{\prod_{t=1}^{|V|=6} P_{2t}^{x_t} (1-P_{2t})^{1-x_t}}_{P(doc|ham=C)}$$

$P(C=ham)\ \pi_h$

Given $x = (x_1, x_2, \ldots, x_{|V|})$

Compare NB1 and NB2.

result = max (NB1, NB2)

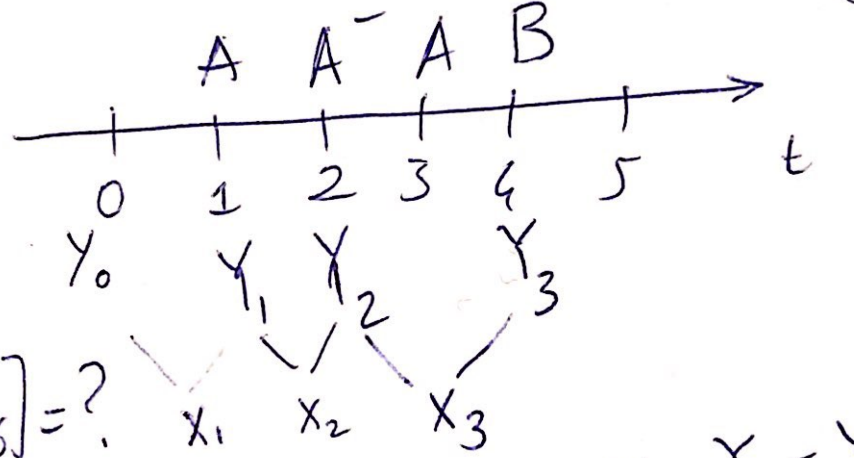
result = max (log NB1, log NB2)

(p_{1t} and p_{2t} are estimated from the training set)

$X = \# K$ – grades $= 6 = k$

How long it will take

$Y_i = \#$ of the ~~new~~ box with a new grade (coupon)



$E[Y_6] = ?$

$X_i = Y_i - Y_{i-1}, \quad X_1 = Y_1 - Y_0$

$$X = Y_6 = \sum_{i=1}^{6} X_i = \sum_{i=1}^{k} X_i$$

$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_5]$$

$$P(X_i) = \frac{k-(i-1)}{k} = \frac{k-(i-1)}{k}, \quad X_i \sim Geo\left(\frac{k-(i-1)}{k}\right)$$

$$E[X_i] = \frac{k}{k-(i-1)}$$ (Geometric Distribution)

$$E[X] = \frac{k}{k} + \frac{k}{k-1} + \cdots + \frac{k}{1} \approx k \log k$$

$$E[X] = \frac{1}{p}, \quad Var = \left(\frac{1-p}{p^2}\right)$$

$$P_r(X = l) = (1-p)^{l-1} p \qquad l = 1, 2, \ldots$$

Probability of success at the lth trial.