tinyurl.com/edmlpr

First tutorial sheet is up

First tutorial next week

Answers released end of next week

Hypothesis Forum

- Share links, code
- Get code review
- Ask questions
- Post answers
  - help others
  - check if right

Generalization

loss function

$$E_{gen} = \mathbb{E}_{p(\underline{x}, y)}\left[L(y, f(\underline{x})) \right]$$

$$\approx \frac{1}{M} \sum_{m=1}^{M} L(y^{(m)}, f(\underline{x}^{(m)})) = E_{test}$$

$\underline{x}^{(m)}, y^{(m)} \sim p(\underline{x}, y)$; M held-out data exampl

---

Need model choice and $\{\underline{x}^{(m)}, y^{(m)}\}$ independen

$\Rightarrow$ Model not chosen using E_{test}

Assume there is a fixed distribution $p(\underline{x}, y)$

How do we avoid fitting test set?

Reduce need to look at test set

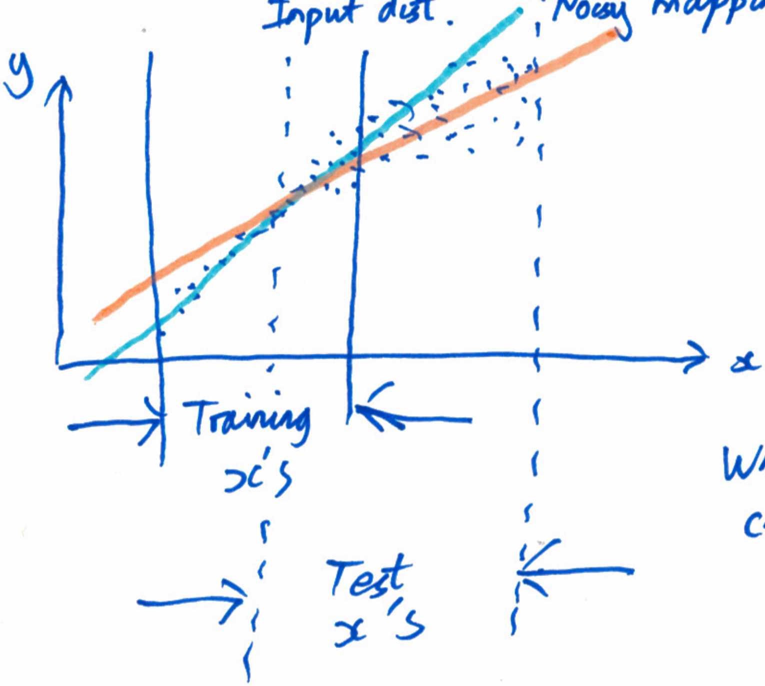
k-fold cross validation

Pick Model (or λ) with lowest average validation SE loss

How do we deal with $p(\underline{x}, y)$ changing?

A/ Depends

$$p(\underline{x}, y) = \underbrace{p(\underline{x})}_{\text{Input dist.}} \underbrace{p(y|\underline{x})}_{\text{Noisy mapping to outputs}} \quad \text{(product rule)}$$



When $p(\underline{x})$ changes.

If $p(y|\underline{x})$ changes? generalization

No y's at ~~test~~ time

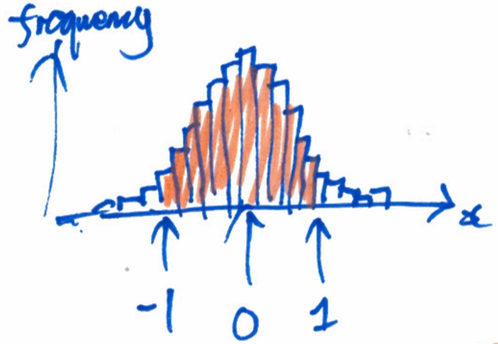
How would we know?

Need some information about the change.

AMOS STORKEY HAS REVIEW

Gaussian (Univariate)

Draw 10^6 values $x \sim N(0,1)$

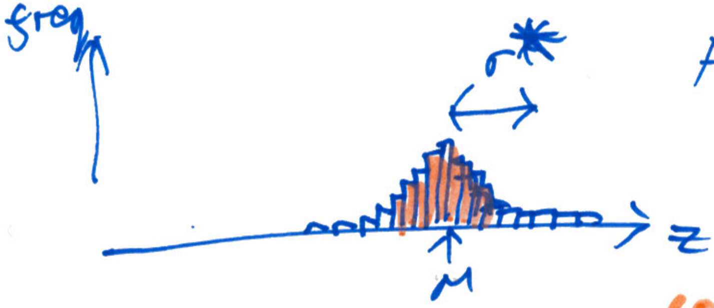


~68% area within ±1

~95% within ±2

$$z = \sigma x + \mu$$

$$x = \frac{z-\mu}{\sigma}$$



Area should be the same.

Height decreased by σ

~68% area within ± σ

$$p(x) = N(x; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$p(z) = N(z; \mu, \sigma^2) = \frac{1}{\underbrace{\sigma\sqrt{2\pi}}_{\sqrt{2\pi\sigma^2}}} e^{-\left(\frac{z-\mu}{\sigma}\right)^2/2}$$

variance

Not every distribution is Gaussian

Probability mass within $\pm\sigma$ $\neq 68\%$ in general

e.g. Load audio file, histogram

on log scale

Central Limit Theorem (CLT)

If x is a sum of or mean
N (many)
independent outcomes, with
finite mean, finite variance

$x \to$ Gaussian as $N \to \infty$

$$x = \frac{1}{N} \sum_{n=1}^{N} q_n$$

Convergence is convergence by distribution

Don't trust Gaussian fit in the tails.

Error bars

$$E_{test} = \frac{1}{M} \sum_m L_m$$

loss on mth example

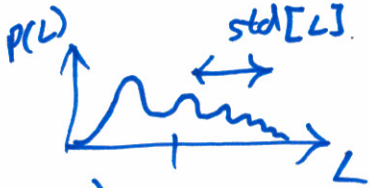
$$\mathbb{E}[E_{test}] = E_{gen}$$

$$\mathrm{var}[E_{test}] = \frac{1}{M^2} \sum_m \mathrm{var}[L_m]$$

(test cases independent)

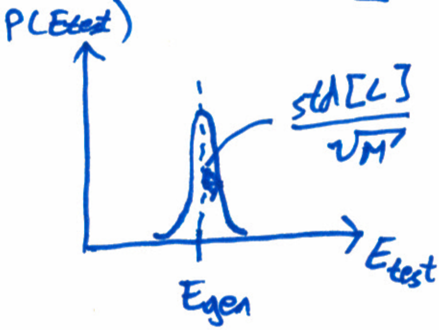
$$= \frac{1}{M^2} \sum_m \mathrm{var}[L]$$

$$= \frac{\mathrm{var}[L]}{M}$$



$$\mathrm{std}[E_{test}] = \sqrt{\frac{\mathrm{var}[L]}{M}}$$

$$= \frac{\mathrm{std}[L]}{\sqrt{M}}$$



$$E_{gen} = E_{test} \pm \frac{\mathrm{std}[L]}{\sqrt{M}}$$