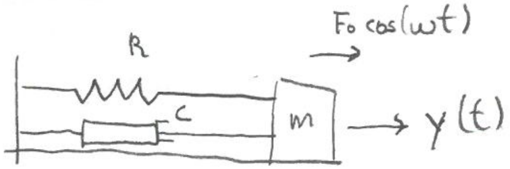
## Three Last Models:



**Model:** $m\dfrac{d^2y}{dt^2} + c\dfrac{dy}{dt} + Ry = F_0 \cos\omega t$

$$y(0) = y_0, \quad \frac{dy}{dt}(0) = y_1$$

**Solution:** $y(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$

$$+ \frac{F_0}{\sqrt{m^2(\omega_0^2 - \omega^2) + c^2\omega^2}\cos(\omega t - \delta)}$$

$$\uparrow\left\{\omega_0^2 = \frac{R}{m}\right.$$

**Special Case:** $y(0) = 1$

$$\frac{dy}{dt}(0) = c = F_0 = 0$$

Then $y(t) = \cos\left(\sqrt{\frac{R}{m}}\, t\right)$

Here

$\theta = (k, m)$

**General:** $\theta = (k, c, m, y_0, y_1)$

**Question:** Which are most uncertain?

**Note:** Take $z_1(t) = y(t)$ and $z_2(t) = \frac{dy}{dt}(t)$ to get

$$\frac{d}{dt}\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -\frac{R}{m} & -\frac{c}{m} \end{bmatrix}\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{F_0 \cos(\omega t)}{m} \end{bmatrix}$$

$$\Rightarrow \frac{dz}{dt} = Az(t) + F(t)$$

$$z(0) = z_0$$

**Model 2:** Scalar exponential process

$$\left.\begin{aligned} \frac{dz}{dt} &= az + b(t) \\ z(0) &= z_0 \end{aligned}\right\} \text{Deterministic}$$

**Inputs:** $\theta = [a, z_0, b(t)]$

**Deterministic Solution:**

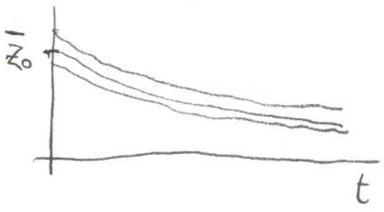
$$z(t, \theta) = e^{at}\left[z_0 + \int_0^t e^{-as} b(s)\right] ds$$

**Random Differential Equation:**

Random variable

$$\frac{dz}{dt} = \overset{\sim}{a(\omega)} z(t) + b(t, \omega)$$

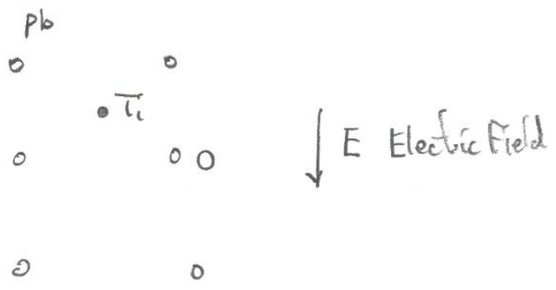
$$z(0) = z_0(\omega)$$

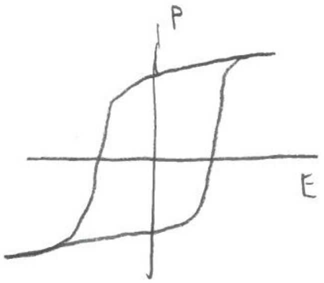


$$\Rightarrow z(t, \omega) = e^{a(\omega)t}\left[z_0(\omega) + \int_0^t e^{-a(\omega)} b(s, \omega) ds\right]$$

[Random field

Model 3: PZT (Lead Zirconate Titanate)

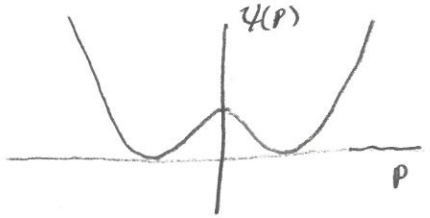




Gibbs Energy: $G(E, P) = \Psi(P) - EP$

Helmholtz Energy: $\Psi(P) = \alpha_1 P^2 + \alpha_{11} P^4 + \alpha_{111} P^6$

$\alpha_1 < 0 \,,\ \alpha_{11} > 0$



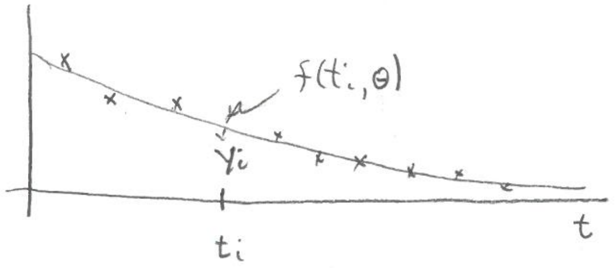
Note: Take $\theta = [\alpha_1, \alpha_{11}, \alpha_{111}]$ to be random variables

Mathematical Model: Describes physical or biological process

Statistical Model: Describes observation process

Example: Take $f(t,\theta) = z(t,\theta) = e^{at}\left[z_0 + \int_0^t e^{-as} b(s)\,ds\right]$

- $f(t_i,\theta)$ : Model predictions at times $t_i$
- $y_i$ : Measured data with errors $\varepsilon_i$

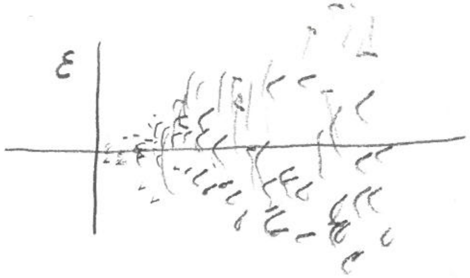


Common Assumptions: Chapter 4

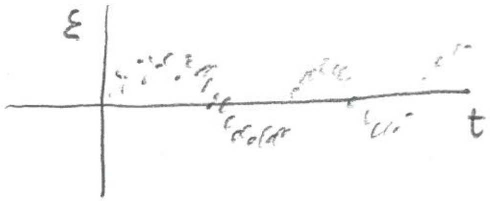
1) $\varepsilon_i$ are independent and identically distributed (iid)

e.g.

iid



Independent but not identically distributed



Not independent

2) $\varepsilon_i \sim N(0, \sigma^2)$

3) $\varepsilon_i \sim P(\lambda)$

Note: Motivates probability and statistics!!

## Common Distributions:

Read Chapter 4

1) Normal (Gaussian)

$X \sim N(\mu, \sigma^2)$ Two parameters: $\mu, \sigma^2$

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

2) $X \sim \mathcal{U}(-1,1)$ Uniform

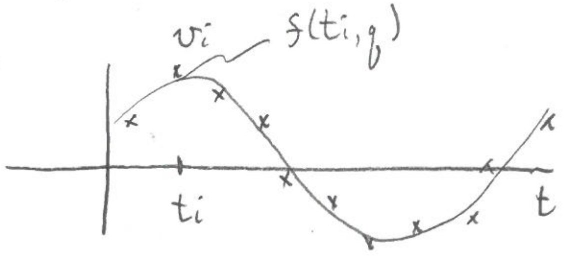
$$f(x) = \frac{1}{2}$$

3) Chi-Squared

$X \sim N(0,1)$

$Y = X^2 \sim \chi^2(1)$ ← 1 degree of freedom

**Motivation:**



**Ordinary Least Squares:** Find $\theta \in \Theta$ that minimizes

$$J(\theta) = \sum_{i=1}^{N} [y_i - f(t_i,\theta)]^2$$

$$\Rightarrow \theta_{min} = \underset{\theta}{\arg\min}\, J(\theta)$$

**Note:**

$$\min_{\theta} J(\theta) = \min_{\theta} \sum_{i=1}^{n} [\varepsilon_i]^2 \quad , \quad \varepsilon_i \sim N(0,\sigma^2)$$

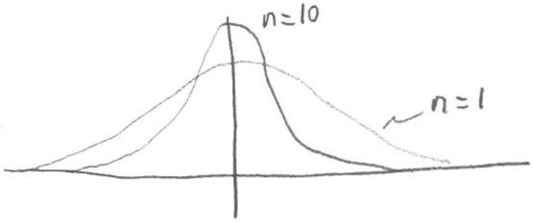
4) Student's T-Distribution

Take $X \sim N(0,1)$ and $Y = X^2 \sim \chi^2(1)$. Suppose $Y_i$ are independent $\chi^2(1)$ and $Z = \sum_{i=1}^{n} Y_i \sim \chi^2(n)$. Then

$$T = \frac{X}{\sqrt{Z/n}}$$

has a student's T-distribution w/ n dof.

**Note:** Used to construct confidence intervals.



**Multiple Random Variables:**

e.g., Polarization model $4(P) = \alpha_{i} P^2 + \alpha_{ii} P^4 + \alpha_{iii} P^6$

$$\theta = [\alpha_{i}, \alpha_{ii}, \alpha_{iii}]$$

**Note:** Let $X = [X, Y]$ be bivariate random variable with joint pdf $f(x,y)$.

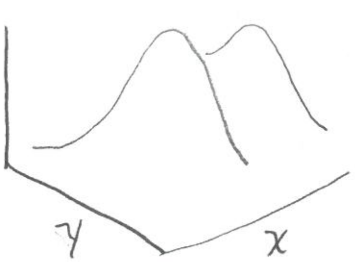
**Marginal:** $f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy$

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx$$

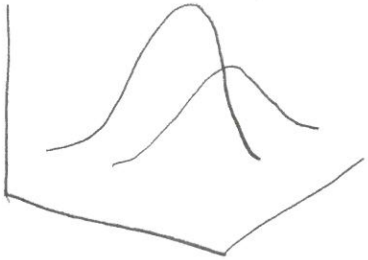
Then $X, Y$ independent if $f(x,y) = f_X(x) f_Y(y)$

**Conditional Density:**

$$f(x|y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & , f_Y(y) > 0 \\ 0 & , \text{else} \end{cases}$$



Marginal



Conditional

**Covariance of X, Y:** e.g., height + weight

$$cov(X,Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$$

$$= \mathbb{E}[XY] - \mathbb{E}(X)\mathbb{E}(Y)$$

**Pearson (Linear) Correlation:**

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \underbrace{\sigma_Y}_{\sqrt{\text{var of } Y}}}$$

**Multivariate Normal:** $X \sim MVN(\mu, V)$ , $X = [X_1, \dots, X_p]$

$$f(X) = \frac{1}{\sqrt{(2\pi)^p \det(V)}} \exp\left[-\frac{1}{2}(X-\mu) V^{-1} (X-\mu)^T\right]$$

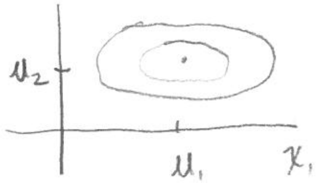
where

$$\mu = [\mu_1, \dots, \mu_p]$$

$$V = cov(X) = \begin{bmatrix} var(X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_p) \\ cov(X_2, X_1) & var(X_2) & & \\ \vdots & & & \\ cov(X_p, X_1) & \cdots & & var(X_p) \end{bmatrix}$$

**Note:** Symmetric, positive definite

**e.g.,** $p = 2$, $cov(X_1, X_2) = 0$



**Note:** $X, Y$ independent $\Rightarrow$ $X, Y$ uncorrelated

Gaussian has $\Leftrightarrow$

**Definition:** An *estimate* ↙ real-valued is a rule or procedure for determining attributes of a quantity based on data.

**Definition:** An *estimator* is associated random variable or random vector.

**e.g.,)** Consider $X_1, \ldots, X_n$. **Goal:** estimate mean & variance

**Estimators:** $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (R.V.) sample mean

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \quad \text{(R.V.) sample variance}$$

**Estimates:** $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ , $s^2$ Similar

**Distributions for the Estimators:**

Suppose $X_i \sim N(\mu, \sigma^2)$. Sampling distributions –

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$S^2 \sim \frac{\sigma^2}{n-1} \chi^2(n-1)$$

$$\Rightarrow \mathbb{E}[S^2] = \sigma^2$$

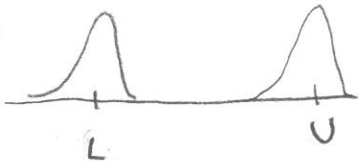
$$\operatorname{var}[S^2] = \left(\frac{\sigma^2}{n-1}\right)^2 \cdot 2(n-1) = \frac{2\sigma^4}{n-1}$$

**Note:** $\chi^2(k)$ ; mean $= k$

variance $= 2k$

Not Symmetric

**Interval Estimators:**



- Exam scores
- Temperatures (Boiling & freezing)

**Goal:** Determine functions $L(x)$ and $U(x)$ that bound the location of $\theta$,

$$L(x) < \theta < U(x),$$

based on realizations $x = [x_1, \ldots, x_n]$ of a random sample $X = [X_1, \ldots, X_n]$.

**Interval Estimator:** Random interval $[L(X), U(X)]$

**Confidence Interval:** Interval estimator plus confidence coef.

- $(1-\alpha) \times 100\%$ Confidence Interval $= [L(X), U(X)]$

such that

$$P[L(X) \leq \theta < U(X)] = 1 - \alpha$$

**Example:** Suppose $X_i \sim N(\mu, \sigma^2)$ where $\mu$ is unknown and $\sigma^2$ is known.

Example 4.32

$x \sim N(\mu, \sigma^2)$

```
>> x = mu + sigma/√n*randn
```

Consider

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

Note:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X}-\mu}{\underbrace{\boxed{\sigma/\sqrt{n}}}_{\sigma_{\bar{X}}}} \sim N(0,1)$$

Then

$$P\left(-2 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 2\right) = 0.9545$$

$$\Rightarrow P\left(\bar{X} - \frac{2\sigma}{\sqrt{n}} < \mu < \bar{X} + \frac{2\sigma}{\sqrt{n}}\right) = 0.9545$$

**Interval Estimator:** $\left[\bar{X} - \frac{2\sigma}{\sqrt{n}}, \bar{X} + \frac{2\sigma}{\sqrt{n}}\right]$

**Note:** 95.45% Confidence Interval: $\left[\bar{x} - \frac{2\sigma}{\sqrt{n}}, \bar{x} + \frac{2\sigma}{\sqrt{n}}\right]$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \leftarrow$ Realization

**Example 4.33:** $\mu, \sigma^2$ both unknown – Need for linear regression

**Note:**

$$X = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$$

$$Z = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

Then

$$T = \frac{X}{\sqrt{Z/(n-1)}} = \frac{\sqrt{n}(\bar{X}-\mu)}{S}$$

has t-distribution with $n-1$ d.o.f.

**Goal:** Find a & b such that

$$P\left(a < \frac{\sqrt{n}(\bar{X}-\mu)}{S} < b\right) = 1-\alpha$$

Symmetry: Take $b = -a$

Notation: $t_{n-1,1-\frac{\alpha}{2}} \Rightarrow n-1$ dof, prob $1-\frac{\alpha}{2}$

Then $a = t_{n-1,1-\frac{\alpha}{2}}$ so

$$P\left(\bar{X} - \frac{t_{n-1,1-\frac{\alpha}{2}} S}{\sqrt{n}} < \mu < \bar{X} + \frac{t_{n-1,1-\alpha/2} S}{\sqrt{n}}\right) = 1-\alpha$$

**Interval:** Employ realizations

$$\left[\bar{x} - \frac{t_{n-1,1-\alpha/2}\, s}{\sqrt{n}}, \bar{x} + \frac{t_{n-1,1-\alpha/2}\, s}{\sqrt{n}}\right].$$

---

**Example:** Consider n height measurements $x_i$ from population with

$\mu = 67$

$\sigma = 2.5$

**Note:** height_example.m

**Assumption:** $X_i \sim N(\mu, \sigma^2)$

MATLAB: >> $X_i = \overset{*\text{ones}(1,n)}{\mu} + \sigma * \text{randn}(1,n)$

Recall: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Then

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$S^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$$

Code: Try with $n = 44$, 400, 4000 and show decrease in variance

Question: How do we plot samples $x_i$?

1.) Histogram: histogram(X, nbin)
2.) Normalized histogram: histnorm(X) ] — Binning is an issue - Demonstrate
3.) Kernel density estimate (kde) • Pages 75-76
 - ksdensity.m MATLAB Statistics toolbox
 - kde.m, kde2d.m MATLAB central

Note: histnorm(X, nbins, 'plot')

Note: Compare interval estimates

## Ordinary Least Squares and Likelihood Estimators: Section 4.3

Recall Statistical Model:

$$Y_i = f(t_i,\theta) + \varepsilon_i \quad , \quad i = 1,\ldots,n$$

$Y_i$ : Random observations with realizations $y_i$

$\varepsilon_i$ : Random observation errors w/ realizations $\epsilon_i$

$\theta$ : Frequentist theory : True but unknown parameters; not r.v.

Bayesian: Random variables w/ distributions

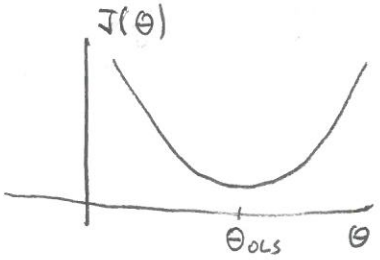
Note: $\theta \in \Theta$ admissible parameter space

OLS Estimator and Estimate:

$$\hat{\Theta}_{OLS} = \underset{\theta \in \Theta}{\text{argmin}} \sum_{i=1}^{n}\left[Y_i - f(t_i,\theta)\right]^2$$

$$\theta_{OLS} = \underset{\theta \in \Theta}{\text{argmin}} \sum_{i=1}^{n}\left[y_i - f(t_i,\theta)\right]^2$$

Typical Assumption: $\varepsilon_i$ iid with true but unknown variance $\sigma^2$ and $\mathbb{E}(\varepsilon_i) = 0$.



## Likelihood Function: Section 4.3.1

Define $f_Y(y,|\theta)$ as parameter-dependent joint pdf for random sample $Y = [Y_1, \dots, Y_n]$.

Note: $\theta = [\theta_1, \dots, \theta_p] \in \Theta$ is unknown parameter

Define $L : \Theta \to [0,\infty)$ by

$$L_y(\theta) = L(\theta|y) = f_Y(y|\theta)$$

where

$\theta$ varies over $\Theta$

$y$ values are fixed

Example: Binomial distribution w/ probability of success $\theta$

$$f_Y(y|\theta,n) = P(Y=y|n,\theta)$$

$$= \binom{n}{y}\theta^{y}(1-\theta)^{n-y} \qquad \text{Discrete}$$

Note: i) Quantifies probability of obtaining exactly $y = 0, 1, \dots, n$ successes in $n$ experiments

ii) $\theta, n$ are known and $y$ is unknown

Likelihood: $L(\theta|y,n) = \binom{n}{y}\theta^{n}(1-\theta)^{n-y}$ continuous

Definition: For $n$ iid random variables, independence yields

$$L(\theta|y) = \prod_{i=1}^{n} f_{Y_i}(y_i,\theta)$$

$$= f_{Y_1}(y_1,\theta)\cdots f_{Y_n}(y_n,\theta)$$

Assumption: $\varepsilon_i \overset{iid}{\sim} N(0,\sigma^2) \Rightarrow Y_i \sim N(f(t_i,\theta),\sigma^2)$

Then

$$L(\theta,\sigma^2|y) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-[y_i - f(t_i,\theta)]^2/2\sigma^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\sum_{i=1}^{n}[y_i - f(t_i,\theta)]^2/2\sigma^2\right]$$

Maximum Likelihood Estimate: For $\theta, \sigma^2$

$$[\theta,\sigma^2]_{MLE} = \underset{\substack{\theta\in\Theta \\ \sigma^2\in(0,\infty)}}{\arg\max}\; L(\theta,\sigma^2|y)$$

Log-Likelihood:

$$l(\theta,\sigma^2|y) = \frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(t_i,\theta))^2$$

Note: $\nabla_\theta l(\theta,\sigma^2|y) = 0$ implies that

$$\sum_{i=1}^{n}[y_i - f(t_i,\theta)]\,\nabla f(t_i,\theta) = 0$$