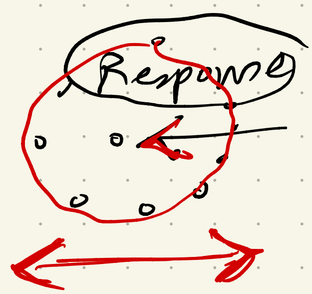
$X \mid \theta \sim Bin(m, \underline{\theta})$ $\quad$ $m = \#$ of items in Box

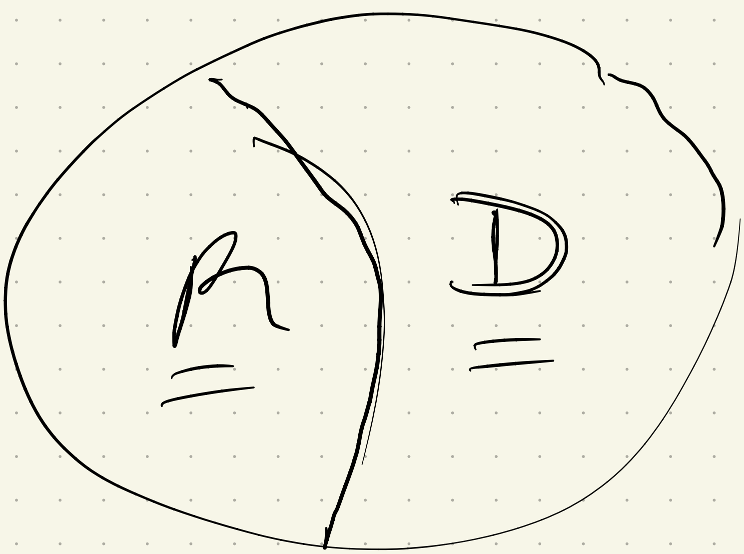
Statistically:

Want to estimate θ

Frequentist paradigm:

$$\hat{\theta} = \frac{X}{m} = \frac{\#\text{ of defectives in Box}}{\text{Size of Box}}$$





Bayesian

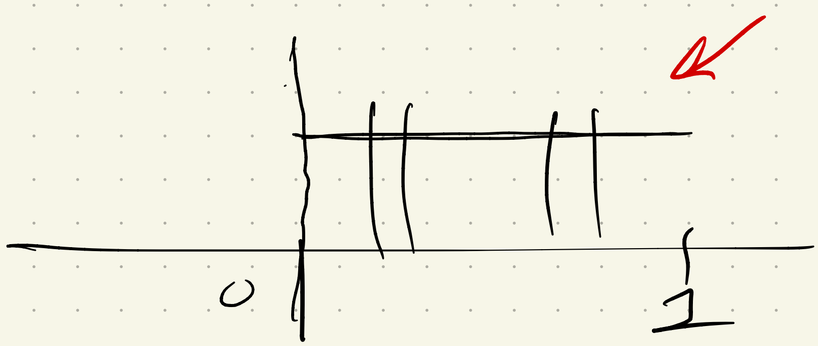
I do know that for a decent company

θ should be small.

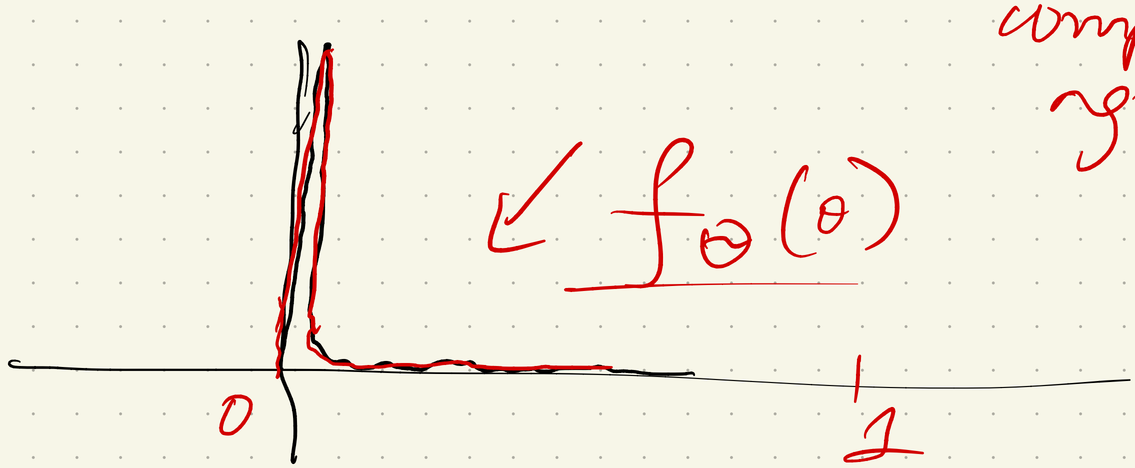
Incorporate this 'prior belief' into your decision making.

THINK about a r.v. $\underline{\underline{\Theta}}$ with some distribution on $(0,1)$ that reflects prior knowledge or opinion.

$$f_{\Theta}(\theta) = \mathbb{1}\{0 < \theta < 1\}$$



you are 'clueless' about where Θ might be completely agnostic.



$f_{\Theta}(\theta)$: putting down this prior

$X \mid \Theta \sim Bin(m, \Theta)$

Decision making depends on the distribution of $\Theta \mid \underline{X}$: how do you update your belief of Θ (starting from the prior) AFTER SEEING the data.

$\to$ posterior of $\Theta \mid X$

Falls exactly into our joint mixed distribution scenario.

Last problem on HW4:

$\Theta \sim Unif(0,1)$

$S_n \mid \Theta \sim Bin(n, \Theta)$

$\Theta \mid S_n \longrightarrow$ Beta distribution

Bayes estimate of Θ is precisely $E[\Theta \mid S_n]$

Example 2:

(Θ, N) : pair of independent random variables, $\Theta \sim \text{Unif}(0,1)$, $N \sim \text{Poisson}(\lambda)$.

$H \mid (N, \Theta) = (n, \theta) \sim \text{Bin}(n, \theta)$

$$f(\theta, n, h)$$

$$= f_{\Theta, N}(\theta, n) \, f_{H \mid \Theta, N}(\theta, n)$$

$$= \frac{e^{-\lambda} \lambda^n}{n!} \binom{n}{h} \theta^h (1-\theta)^{n-h} \, \mathbb{1}\{0 < \theta < 1\} \, \mathbb{1}\{0 \leq h \leq n\}$$

Compute: $f_H(h)$

$$\boxed{f_H(h) = \int_{[0,1]} \left(\sum_{n=h}^{\infty} f(\theta, n, h) \right) d\theta}$$

$$= \frac{\lambda^h}{h!} \int_{[0,1]} e^{-\lambda\theta} \theta^h \, d\theta$$

EH, Var H

$=$

Change of Variable Theorem for Multiple dimensions:

Theorem 4.6.12 of CB.

$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$: random vector with density $f_{\underline{X}}(x_1, \dots, x_n)$

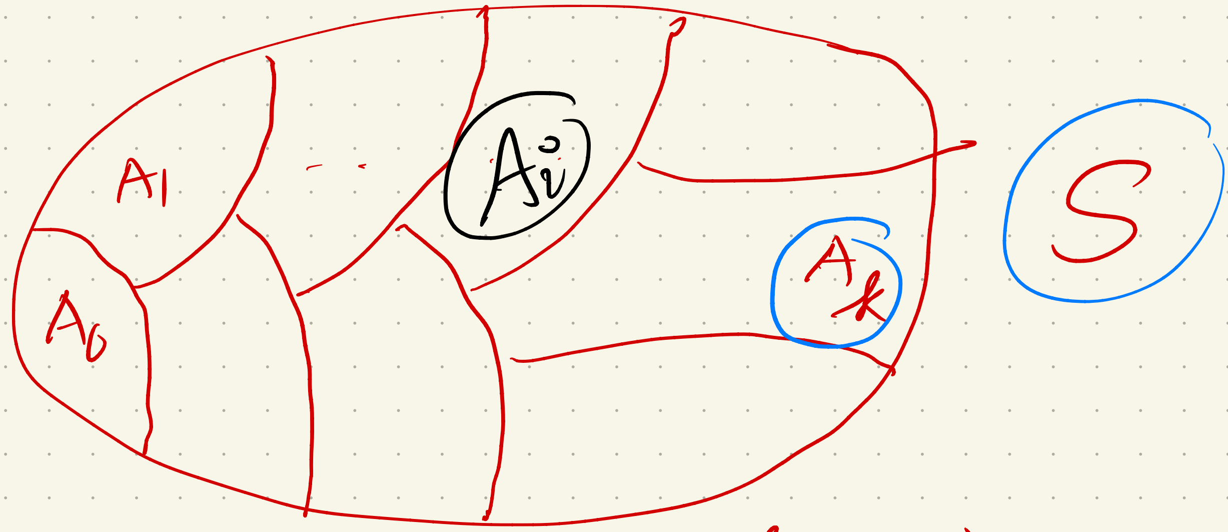
continuous r.v.

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = g\left(\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}\right)$$

$g : \mathbb{R}^n \to \mathbb{R}^n$

$$= \begin{pmatrix} g_1(x_1, \dots, x_n) \\ \vdots \\ g_n(x_1, \dots, x_n) \end{pmatrix}$$

each $g_i : \mathbb{R}^n \to \mathbb{R}$



$P[g(X_1, \dots, X_n) \in S] = 1$

$A_0 \cup A_1 \dots \cup A_k$: partition of S.

$P[(X_1, \dots, X_n) \in \underline{A_0}] = \underline{0}$

$\underline{g|_{A_i}}$ is a one-one and onto transformation from each $A_i \longrightarrow g(S)$

which is <u>continuously</u> <u>differentiable</u>.

h_i: $g(S) \longrightarrow A_i$ which is continuously differentiable, is the inverse function of $g|_{A_i}$

Suppose that $\nabla g(x_1, \dots, x_n)$ is NON-SINGULAR apart from <u>finitely</u> many points

$$\underleftrightarrow{\nabla g(x_1, \ldots, x_n)} = \left(\left(\frac{\partial g_\ell}{\partial x_j}\right)\right)_{\substack{1 \le \ell \le n \\ 1 \le j \le n}}$$

$$= \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}$$

non-singularity

$$\underline{J_g(x_1, \ldots, x_n)} = \det \nabla g(x_1, \ldots, x_n) \neq 0.$$

THEN:

$h_i : g(S) \longrightarrow A_i$

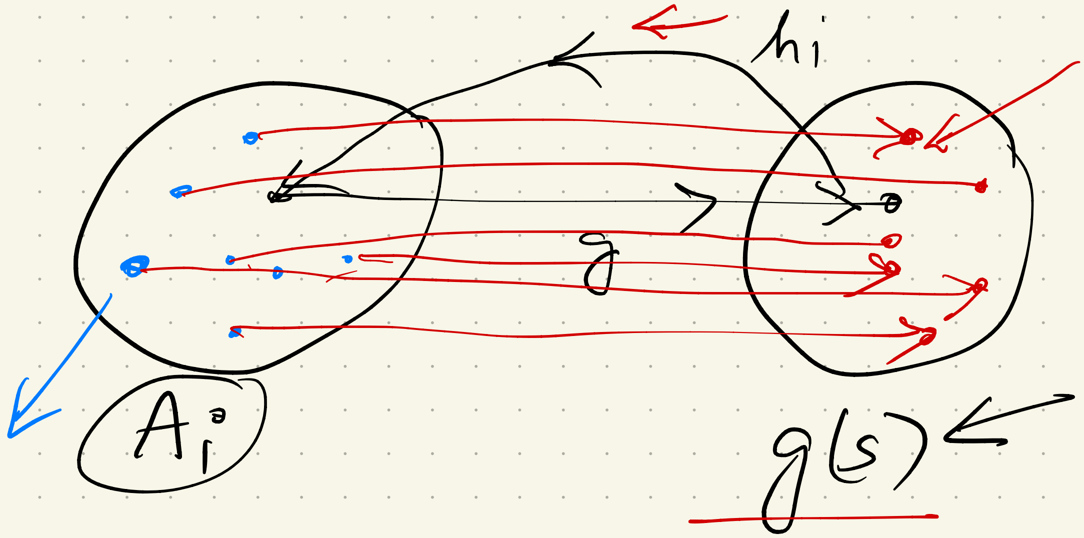
that is also 1-1, onto, continuously differentiable and

$$\underline{J_{h_i}(y_1, \ldots, y_n)} = \det\left[\nabla h_i(y_1, \ldots, y_n)\right] \neq 0$$

$(y_1, \ldots, y_n) \in g(S)$

$$\underline{J_{h_i}}(\underline{y_1, \ldots, y_n}) = \left\{ \left[J_g(\underleftrightarrow{x_1, \ldots, x_n}) \right]^{-1} \right\}$$

where $\underline{g(x_1, \ldots, x_n)} = \underline{(y_1, \ldots, y_n)}$



Bad points such that $J_g(x_1, \ldots, x_n) = 0$

Building up towards writing down an expression for the joint density of $(Y_1, \ldots, Y_n) = \underline{Y}$

$$f_{\underline{Y}}(y_1, \ldots, y_n) = \sum_{i=1}^{k} \boxed{f_X(h_i(y_1, \ldots, y_n))} \, \left| J_{h_i}(y_1, \ldots, y_n) \right|$$

Form is quite similar to the COVT for 1 dimension.

$$J_{h_i}(y_1, \ldots, y_n)$$

$$= \left(\left(\frac{\partial h_{i\ell}}{\partial y_j}(y_1, \ldots, y_n) \right) \right)_{\substack{1 \le \ell \le n \\ 1 \le j \le n}}$$

(X_1, X_2) i.i.d $\mathrm{Exp}(\lambda)$ random variables.

$$f_{X_1, X_2}(x_1, x_2) = \lambda e^{-\lambda x_1} \mathbb{1}(x_1 > 0) \, \lambda e^{-\lambda x_2} \mathbb{1}(x_2 > 0)$$

$$= \lambda^2 e^{-\lambda (x_1 + x_2)} \mathbb{1}(x_1 > 0, x_2 > 0)$$

$$\left. \begin{aligned} Y_1 &= \frac{X_1 + X_2}{X_1} \\ Y_2 &= \frac{X_1}{X_1 + X_2} \end{aligned} \right\} \begin{aligned} &= \underline{g_1(X_1, X_2)} \\ &= g_2(X_1, X_2) \end{aligned}$$

$$g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} : \underline{\mathbb{R}^2} \to \underline{\mathbb{R}^2}$$

$$g : \underline{\mathbb{R}_+^2}_{S} \longrightarrow \mathbb{R}_+ \times (0,1) = \underline{g(S)}$$

$$g(x_1, x_2) = \left(x_1 + x_2, \frac{x_1}{x_1 + x_2} \right)$$

$$= (y_1, y_2)$$

$$\mathbb{R}_+ = [0, \infty)$$

g : is a 1-1 onto map between S and $g(S)$.

$$\underline{y_1 = x_1 + x_2}$$

$$\underline{y_2 = \frac{x_1}{x_1 + x_2}} \Rightarrow y_2(x_1 + x_2) = x_1.$$

Therefore: $\cancel{y_1 =}$ $x_1 = y_2(\underline{x_1 + x_2}) = y_2 y_1$

$$x_2 = y_1 - x_1$$

$$= y_1(1 - y_2)$$

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(y_1 y_2,\ y_1(1-y_2))\,y_1$$

$$\underline{f_{X_1,X_2}(x_1,x_2) = \lambda^2 e^{-\lambda(x_1+x_2)}}$$

$$f_{Y_1,Y_2}(y_1,y_2) = \lambda^2 e^{-\lambda(y_1 y_2 + y_1(1-y_2))}\,y_1\ \mathbb{1}\{y_1>0,\ 0<y_2<1\}$$

$$= \lambda^2 e^{-\lambda y_1}\,y_1\,\underline{\mathbb{1}\{y_1>0,\ 0<y_2<1\}}$$

This gives the joint density of the transformed vector (Y_1,Y_2)

$$= \lambda^2 e^{-\lambda y_1}\,y_1\,\mathbb{1}(y_1>0)\,\boxed{\mathbb{1}(0<y_2<1)}$$

Notice that the joint density factor as $\psi_1(y_1)\,\psi_2(y_2)$ implying ~~that~~ Y_1 and Y_2 are independent

$Y_2 \sim \text{Unif}(0,1)$

$Y_1 \sim \text{Gamma}(2, \lambda) \longleftarrow \frac{1}{}$

am NOT using the CB parametrization.

Example 4.6.13: Take a look.

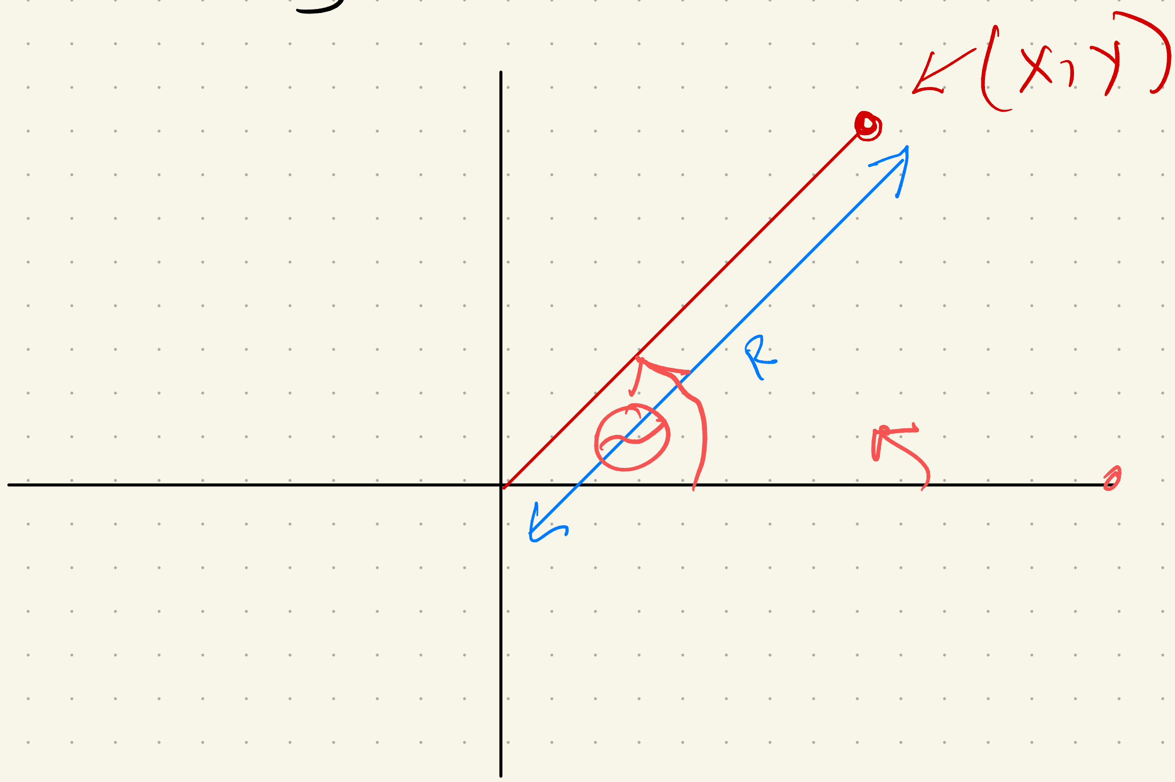
Exercise: $\left.\begin{array}{l} X \sim N(0,1) \\ Y \sim N(0,1) \end{array}\right\}$.

X and Y are independent.

Let $X = R\cos\Theta$, $\quad 0 < R < \infty$

$Y = R\sin\Theta$ $\quad 0 < \Theta < 2\pi$

Use the COVT to find the joint density of (R, Θ).



Hint: I've actually already written out the inverse transformation