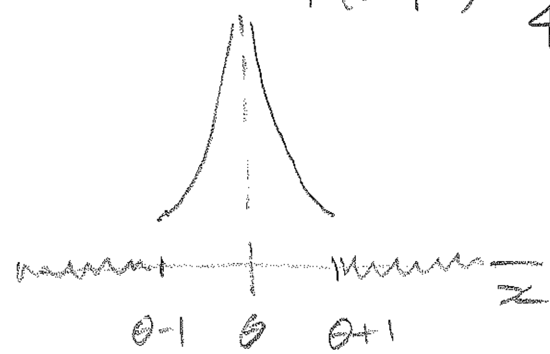
<u>Example</u> Since $\int_0^1 x^{-\frac{1}{2}}\,dx = 2x^{\frac{1}{2}}\Big|_0^1 = 2$

the following is a density

$$f(x|\theta) = \frac{1}{4|x-\theta|^{1/2}}\, I\left[|x-\theta|<1\right]$$



But for $X_1, \ldots, X_n$ iid $f(x|\theta)$, the joint density is

$$f(x|\theta) = \prod_{i=1}^{n} \frac{1}{4|x_i-\theta|^{1/2}}\, \underbrace{I\left[\text{all } |x_i-\theta|<1\right]}_{I\left[\,x_{(n)}-1<\theta<x_{(1)}+1\right]}$$

But if there is any $x_i \in (x_{(n)}-1,\ x_{(1)}+1)$ $L(\theta)$ is "infinite" at such x_i — i.e. it's possible to have even multiple singularities

<u>Example</u> $X \sim N(\mu, \sigma^2)$ for parameter (μ, σ^2)

$$L(\mu, \sigma^2) = f(x|\mu, \sigma^2)$$

and $L(X, \sigma^2) \longrightarrow \infty$ as $\sigma^2 \longrightarrow 0$ i.e. $L(\mu, \sigma^2)$ is unbounded on $\mathbb{R} \times (0, \infty)$

One way of addressing the possibility of an unbounded likelihood (a favorite of Prof. Meeker, e.g.) is to admit that all data are really discrete and measured to the nearest something (say Δ) — a cont's model is really an idealization for "data" that have infinite # of decimal places — i.e. for a density $f(x|\theta)$ I could use probabilities

$$f^*(c|\theta) = P_\theta\left[c - \frac{\Delta}{2} < X < c + \frac{\Delta}{2}\right] = \int_{c-\frac{\Delta}{2}}^{c+\frac{\Delta}{2}} f(x|\theta)\,dx$$

to form a (discrete data) likelihood that is of necessity ≤ 1 (it's a joint probability)

<u>Example</u> Ideal (cont's) observations $X_1, \ldots, X_n$ iid $N(\mu, \sigma^2)$ — Suppose

$Y_i = X_i$ rounded to nearest integer

& discrete with pmf

$$f^*(y|\mu,\sigma^2) = \Phi\left(\frac{y+.5-\mu}{\sigma}\right) - \Phi\left(\frac{y-.5-\mu}{\sigma}\right)$$

So that the likelihood is

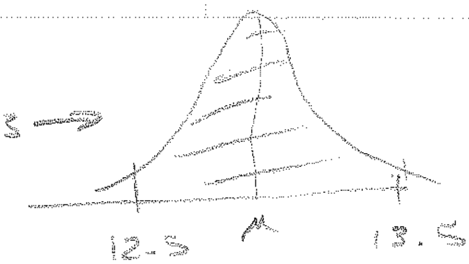
$$L(\mu,\sigma^2) = \prod_{i=1}^{n}\left[\Phi\left(\frac{y_i+.5-\mu}{\sigma}\right) - \Phi\left(\frac{y_i-.5-\mu}{\sigma}\right)\right] \leq 1$$

If all $y_i = 13$, then

$$L(\mu,\sigma^2) = \left[\Phi\left(\frac{13.5-\mu}{\sigma}\right) - \Phi\left(\frac{12.5-\mu}{\sigma}\right)\right]^n$$

Of course $L(\mu,\sigma^2) < 1$, but by taking μ to be any element of $(12.5, 13.15)$ since

$f^*(13|\mu,\sigma^2)$ is this $\Rightarrow$



I can send $L(\mu,\sigma^2) \to 1$ by sending $\sigma^2 \to 0$ – the supremum of $L(\mu,\sigma^2)$ is never achieved – so even though the infinity problem is eliminated there is no maximizer for $(\mu,\sigma^2) \in \mathbb{R} \times (0,\infty)$

<u>Example</u> The $U(\theta_1,\theta_2)$ example I began with strictly speaking has no MLE if I use a density positive on the <u>open</u> interval (θ_1,θ_2) in the problem set-up (instead of $f(x|\theta_1,\theta_2) = \frac{1}{\theta_2-\theta_1} I[\theta_1 \le x \le \theta_2]$) strictly speaking, there is no maximizer of the likelihood –

Both in terms of computations employed to find MLE's in practice and in terms of the theory used to describe properties of maximum likelihood, the topic is often approached somewhat indirectly as follows

Suppose $\Theta \subset \mathbb{R}^k$ – If Θ is open and $L(\theta)$ is differentiable, a <u>necessary</u> condition for θ^* to maximize $L(\theta)$ is that all first partials of

$$l(\theta) = \log L(\theta)$$

be 0, i.e. That

$$\left.\frac{\partial}{\partial \theta_1} \ell(\theta)\right|_{\theta=\theta^*} = 0$$

$$\vdots$$

$$\left.\frac{\partial}{\partial \theta_k} \ell(\theta)\right|_{\theta=\theta^*} = 0$$

These are the "likelihood" equations

i.e. $\underbrace{\nabla \ell(\theta^*)}_{\text{score function}} = \underset{\text{vector } 0}{0}$ — The "estimating equation" for ML

Why the log? maximizing $\ell(\theta)$ is equivalent to maximizing $L(\theta)$ and for iid cases the log turns a product into a sum, which is more tractable —

day 16

Example $X \sim$ Binomial $(5, p)$ — pictures from day #1

$$L(p) = \binom{5}{x} p^x (1-p)^{5-x}$$

$$\ell(p) = \log\binom{5}{x} + x \log p + (5-x)\log(1-p)$$

$$\ell'(p) = \frac{x}{p} - \frac{5-x}{1-p}$$ score function

$x \neq 0, 5$ $\quad \ell'(p) = 0 \Rightarrow \frac{x}{p} = \frac{5-x}{1-p}$

$$x - xp = 5p - xp$$

$$p = \frac{x}{5}$$

(and $\ell''(p) = -\frac{x}{p^2} - \frac{5-x}{(1-p)^2} < 0$)

For $x = 0$ $\quad \ell'(p) = -\frac{5}{1-p} < 0$

and $L(p)$ is decreasing