

2.4 Tools

INTERCHANGE INTEGRAL AND DIFFERENTIAL

Theorem (Leibniz's Rule) if

- $f(x, \theta)$, $a(\theta)$, $b(\theta)$ are differentiable with respect to θ
- $\frac{\partial f}{\partial \theta}(x, \theta)$ is continuous on $x \in (x_0, x_1)$ \leftarrow can be weakened

Then for $a(\theta), b(\theta) \in (x_0, x_1)$:

$$\begin{aligned} \frac{d}{d\theta} \int_{a(\theta)}^{b(\theta)} f(x, \theta) dx &= f(b(\theta), \theta) \frac{d}{d\theta} b(\theta) - f(a(\theta), \theta) \frac{d}{d\theta} a(\theta) \\ &\quad + \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\partial \theta} f(x, \theta) dx \end{aligned}$$

REMARK

- if $a(\theta) = a$, $b(\theta) = b$ then

$$\frac{d}{d\theta} \int_a^b f(x, \theta) dx = \int_a^b \frac{\partial f}{\partial \theta}(x, \theta) dx$$

Proof

$$LHS = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \left(\int_{a(\theta+\delta)}^{b(\theta+\delta)} f(x, \theta+\delta) dx - \int_{a(\theta)}^{b(\theta)} f(x, \theta) dx \right)$$

$$\begin{aligned} &= \lim_{\delta \rightarrow 0} \frac{1}{\delta} \left[\int_{a(\theta)}^{b(\theta)} (f(x, \theta+\delta) - f(x, \theta)) dx \right. \\ &\quad \left. + \int_{a(\theta)}^{a(\theta)} f(x, \theta+\delta) dx + \int_{b(\theta)}^{b(\theta+\delta)} f(x, \theta+\delta) dx \right] \end{aligned}$$

$$+ \int_{a(\theta+\delta)}^{\theta} f(x, \theta+\delta) dx + \int_{\theta}^{b(\theta)} f(x, \theta+\delta) dx \\ =: A + B + C$$

where $b(\theta)$ fixed, finite boundaries

$$A = \lim_{\delta \rightarrow 0} \frac{1}{\delta} \int_{a(\theta)}^{b(\theta)} (f(x, \theta+\delta) - f(x, \theta)) dx$$

$$\stackrel{(1)}{=} \int_{a(\theta)}^{b(\theta)} \lim_{\delta \rightarrow 0} \frac{1}{\delta} (f(x, \theta+\delta) - f(x, \theta)) dx$$

$$= \int_{a(\theta)}^{b(\theta)} \frac{\partial f}{\partial \theta}(x, \theta) d\theta \quad (\text{Since } \frac{\partial f}{\partial \theta} \text{ is cont. in } x \text{ and next theorem})$$

$$B = \lim_{\delta \rightarrow 0} -\frac{1}{\delta} \int_{a(\theta)}^{a(\theta+\delta)} f(x, \theta+\delta) dx$$

$$= \lim_{\delta \rightarrow 0} -\frac{a(\theta+\delta) - a(\theta)}{\delta} \times \frac{1}{a(\theta+\delta) - a(\theta)} \int_{a(\theta)}^{a(\theta+\delta)} f(x, \theta+\delta) dx$$

$$= -\frac{d}{d\theta} a(\theta) \times f(a(\theta), \theta) \quad (a(\theta) \text{ is diff.})$$

Likewise

$$C = \frac{d}{d\theta} b(\theta) f(b(\theta), \theta) . \quad \square$$

To strengthen the above result for infinite domain of the integral, we need a new tool

Thm (interchanging limit and integral)

Suppose

- $h(x, y)$ is continuous at $y = y_0$ for each fixed x
- There is a function $g(x)$ such that

envelope condition

$$\left\{ \begin{array}{l} |h(x, y)| \leq g(x) \quad \forall x, y \\ \int_{-\infty}^{\infty} g(x) dx < \infty \end{array} \right.$$

Then

$$\lim_{y \rightarrow y_0} \int_{-\infty}^{\infty} h(x, y) dx = \int_{-\infty}^{\infty} \lim_{y \rightarrow y_0} h(x, y) dx$$

Remark

- g is called the dominating OR envelope function of f .
- Known as (a version / consequence of) Lebesgue's Dominated Convergence theorem (D.C.T.)

INTERCHANGE Differential And integral

Thm. Suppose

- For each x , $f(x, \theta)$ is differentiable wrt θ , at $\theta \in \Theta$
- For each $\theta \in \Theta$, there is a function $g(\cdot, \theta)$ and $\delta_0 > 0$ such that

$$\left\{ \begin{array}{l} \left| \frac{\partial}{\partial \theta} f(x, \theta) \right|_{\theta=\theta'} \leq g(x, \theta') + \delta' \in (\theta - \delta_0, \theta + \delta_0) \\ \int_{-\infty}^{\infty} g(x, \theta) dx < \infty \quad \text{for each } \theta \in \Theta. \end{array} \right.$$

envelope condition

Then

$$\frac{d}{d\theta} \int_{-\infty}^{\infty} f(x, \theta) dx = \int_{-\infty}^{\infty} \frac{d}{d\theta} f(x, \theta) dx.$$

holds for each $\theta \in \Theta$.

Remark.

- Leibniz's Rule holds under this weaker envelope condition for $f(x, \theta)$.
- if we need only to differentiate at $\theta = \theta_0$, then it is sufficient that the envelope condition be satisfied for a neighborhood of θ_0 i.e. $\forall \theta \in (\theta_0 - \delta_0, \theta_0 + \delta_0)$ for some $\delta_0 > 0$, and only a dominating function $g(x, \theta_0)$ (for fixed $\theta = \theta_0$) is required.
- This Thm is a direct consequence of Lebesgue's Dominated convergence Theorem.

Example (Exponential)

$$X \sim \text{Exp}(\lambda)$$

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \quad x \in [0, +\infty)$$

Moments: $\mathbb{E} X^n = \int_0^\infty x^n \frac{1}{\lambda} e^{-x/\lambda} dx, n=1,2,\dots$

Wish to differentiate

$$\frac{d}{d\lambda} \mathbb{E} X^n = \frac{d}{d\lambda} \int_0^\infty$$

$$= \int_0^\infty \frac{d}{d\lambda} \quad (\star) \text{ need to justify this}$$

$$= \int_0^\infty x^n \left(-\frac{1}{\lambda^2} e^{-x/\lambda} + \frac{1}{\lambda} e^{-x/\lambda} \frac{x}{\lambda^2} \right) dx$$

$$= \int_0^\infty x^n e^{-x/\lambda} \frac{1}{\lambda^2} \left(\frac{x}{\lambda} - 1 \right) dx$$

$$= \int_0^\infty x^{n+1} e^{-x/\lambda} \frac{1}{\lambda^3} - x^n e^{-x/\lambda} \frac{1}{\lambda^2} dx$$

$$= \frac{1}{\lambda^2} \mathbb{E} X^{n+1} - \frac{1}{\lambda} \mathbb{E} X^n$$

This gives a Recursion Relation

$$\mathbb{E} X^{n+1} = \lambda \mathbb{E} X^n + \lambda^2 \frac{d}{d\lambda} \mathbb{E} X^n.$$

Similar identities hold for a broad family of distributions

Similar identities hold for a broad family of distributions

justifying the interchange of \int and $\frac{d}{d\lambda}$
the derivative of the integrand is

$$\frac{d}{dx} f(x, \lambda) = x^n e^{-x/\lambda} \frac{1}{\lambda^2} \left(\frac{x}{\lambda} - 1 \right)$$

Need to find a dominating function, $g(\cdot, \lambda)$ s.t

- ① $\left| \frac{d}{d\theta} f(x, \theta) \Big|_{\theta=\theta_0} \right| < g(x, \lambda) \quad \forall \theta' \in (\lambda - \delta_0, \lambda + \delta_0)$
- ② and g integrable wrt x . for some $\delta > 0$

$$\begin{aligned} \left| \frac{d}{d\theta} f(x, \theta) \right| &= x^n e^{-x/\theta} \frac{1}{\theta^2} \left| \frac{x}{\theta} - 1 \right| \\ &\leq x^n e^{-x/\theta} \frac{1}{\theta^2} \left(\frac{x}{\theta} + 1 \right), \quad \text{since } x > 0 \\ &\leq x^n e^{-x/(\theta+\delta_0)} \frac{1}{(\theta-\delta_0)^2} \left(\frac{x}{\theta-\delta_0} + 1 \right) =: g(x, \theta) \end{aligned}$$

The last inequality holds $\forall x > 0, \forall \theta > \delta_0$.

Thus ① holds for the chosen $g(x, \lambda)$.

To verify ②

$$\int g(x, \lambda) dx = \int x^{n+1} e^{-x/\lambda + \delta_0} \frac{1}{(\lambda - \delta_0)^3} + \int x^n e^{-x/\lambda + \delta_0} \frac{1}{(\lambda - \delta_0)^2} dx$$

= multiple of $(n+1)$ -th moment of an exponential Rv
+ multiple of n -th moment of another exp. Rv

+ multiple of n -th moment of another Exp. RV
 $< +\infty$. \square

Example (Gaussian)

$$X \sim N(0,1)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, x \in \mathbb{R}$$

$$\text{MGF} \quad M_X(t) = \mathbb{E} e^{tx} = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

"Generating Moments" via

$$\frac{d}{dt} M_X(t) = \frac{d}{dt} \mathbb{E} e^{tx}$$

$$(*) \quad \textcircled{=} \quad \mathbb{E} \frac{d}{dt} e^{tx} = \mathbb{E} X e^{tx}.$$

Justifying (*) by finding dominating function for

$$\left| \frac{d}{dt} e^{tx} e^{-x^2/2} \right| \\ = |x e^{tx} e^{-x^2/2}| = |x| e^{tx} e^{-x^2/2}$$

Let $0 < s_0 < t$:

$$\text{if } x \geq 0: |x e^{tx} e^{-x^2/2}| \leq x e^{(t+s_0)x} e^{-x^2/2}$$

$$\text{if } x < 0: |x e^{tx} e^{-x^2/2}| \leq (-x) e^{(t-s_0)x} e^{-x^2/2}.$$

Take the RHS as $g(x,t)$ for $x \geq 0$ and $x < 0$.

Then $\left| \frac{d}{dt} e^{tx} e^{-x^2/2} \right| \leq g(x,t) \quad \forall t' \in (t-s_0, t+s_0)$
 The domination holds.

Then, one can check that $g(\cdot, t)$ is integrable \square

INTERCHANGING SUM AND DIFFERENTIAL

Suppose $\sum_{x=0}^{\infty} h(x, \theta)$ exists (i.e. converges pointwise) $\forall \theta \in (a, b)$

Moreover, assume

(1) $\frac{\partial}{\partial \theta} h(x, \theta)$ is continuous in θ for each x

(2) $\sum_{x=0}^{\infty} \frac{\partial}{\partial \theta} h(x, \theta)$ converges uniformly for all θ in a closed subinterval of (a, b)

Then

$$\frac{d}{d\theta} \sum_{x=0}^{\infty} h(x, \theta) = \sum_{x=0}^{\infty} \frac{\partial}{\partial \theta} h(x, \theta)$$

Remark.

This is a consequence of Lebesgue's D.C.T.

Example $X \sim \text{Geometric}(\theta)$, $0 < \theta < 1$

$$P(X=x) = \theta(1-\theta)^x, x=0,1,2,\dots$$

this is a valid pmf since

$$\sum_{x=0}^{\infty} \theta(1-\theta)^x = 1$$

(*) Differentiating both sides wrt θ , assuming we can interchange \sum and $\frac{d}{d\theta}$:

$$\sum_{x=0}^{\infty} (1-\theta)^x - \theta x (1-\theta)^{x-1} = 0$$

$$\Rightarrow \sum_{x=0}^{\infty} (1-\theta)^x = \sum_{x=0}^{\infty} \theta x (1-\theta)^{x-1}$$

$$\underbrace{\frac{1}{1-(1-\theta)}} = \frac{1}{1-\theta} \sum_{x=0}^{\infty} \theta x (1-\theta)^x$$

$$\frac{1}{\theta} = \frac{1}{1-\theta} \mathbb{E}X$$

$$\text{Hence } \mathbb{E}X = \frac{\theta}{1-\theta},$$

justifying (*)

- $h(x, \theta) = \theta(1-\theta)^x$
 $\forall \theta \in (0, 1)$ $\sum_{x=0}^k h(x, \theta) = \theta \sum_{x=0}^k (1-\theta)^x = \theta \frac{1-(1-\theta)^{k+1}}{1-(1-\theta)} \rightarrow 1$ as $k \rightarrow \infty$.
- $\frac{\partial}{\partial \theta} h(x, \theta) = (1-\theta)^x - \theta x (1-\theta)^{x-1}$
 continuous in θ
 So h is continuously diff. in θ .
- "Uniform convergence" of $\sum_{x=0}^k \frac{\partial}{\partial \theta} h(x, \theta)$

Pointwise Convergence: for each $\theta \in (0, 1)$

$$\begin{aligned}
 S_k(\theta) &:= \sum_{x=0}^k \frac{\partial}{\partial \theta} h(x, \theta) = \sum_{x=0}^k (1-\theta)^x - \underbrace{\theta x (1-\theta)^{x-1}}_{} \\
 &= \frac{1-(1-\theta)^{k+1}}{\theta} + \theta \sum_{x=0}^k \frac{\partial}{\partial \theta} (1-\theta)^x \\
 &= \frac{1-(1-\theta)^{k+1}}{\theta} + \theta \frac{d}{d\theta} \sum_{x=0}^k (1-\theta)^x \\
 &= \frac{1-(1-\theta)^{k+1}}{\theta} + \theta \frac{d}{d\theta} \frac{1-(1-\theta)^{k+1}}{\theta} \\
 &= \frac{1-(1-\theta)^{k+1}}{\theta} + \theta \underbrace{\frac{(k+1)(1-\theta)^k}{\theta^2} \theta - (1-(1-\theta)^{k+1})}_{}
 \end{aligned}$$

$$= \frac{1-(1-\theta)}{\theta} + \theta \underbrace{\dots}_{\theta^2}$$

$$= (k+1)(1-\theta)^k$$

So $S_k(\theta) \rightarrow 0$ as $k \rightarrow \infty$.

Uniform Convergence: Take any $[c, d] \in (0, 1)$

check:

$$\sup_{\theta \in [c, d]} |S_k(\theta) - 0| = \sup_{\theta \in [c, d]} (k+1)(1-\theta)^k$$

$$\rightarrow 0 \text{ as } k \rightarrow \infty$$

□

INTERCHANGING INTEGRAL AND SUM

Thm

Suppose

$$\sum_{x=0}^{\infty} h(x, \theta) \text{ exists (i.e. converges pointwise)} \quad \forall \theta \in [a, b]$$

Moreover, assume

① $h(x, \theta)$ is continuous in θ for each fixed x

② $\sum_{x=0}^{\infty} h(x, \theta)$ converges uniformly on $[a, b]$

Then

$$\boxed{\int_a^b \sum_{x=0}^{\infty} h(x, \theta) d\theta = \sum_{x=0}^{\infty} \int_a^b h(x, \theta) d\theta}$$

Remark.

This is also a consequence of Lebesgue's D.C.T.