2.1

- If X is a random variable a function $g(x)$ is also a random variable, say $Y=g(x)$ is a new RV
$$\forall \text{ set } A \quad P(Y\in A) = P(g(x)\in A)$$

So: $g(x): \mathcal{X} \to \mathcal{Y}$ (g maps $\mathcal{X}$ to $\mathcal{Y}$)

- We would also like to associate an inverse mapping, g^{-1}
So: $g^{-1}(y): \mathcal{Y} \to \mathcal{X} \ni g^{-1}(A) = \{x\in\mathcal{X} \mid g(x)\in A\}$
We note: $g^{-1}(A)$ returns the set of x that $g(x)$ takes into A
Now: $P(Y\in A) = P(g(x)\in A) = P(\{x\in X \mid g(x)\in A\}) = P(x\in g^{-1}(A))$
defines the probability distribution of Y

Example: $f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x} \quad \forall\, x=0,1,2,\dots n$
$0\le p\le 1$; $n \ne p$

Consider $Y=g(x)=n-x$ Since the support of X is $0,1,2,\dots n$ the support of y is also $0,1,2,\dots$ not b/c it copies over, but b/c $n-n=0,\ n-(n-1)=1,\ \dots\ n-0=n$

So $\forall\, y\in Y$, $n-x=g(x)=y$ iff $x=n-y\in g^{-1}(y)$

So $f_Y(y) = \sum_{x\in g^{-1}(y)} f_X(x)$ (PDF)
$= f_X(n-y)$
$= \binom{n}{n-y} p^{n-y} (1-p)^{n-n-y}$
$= \binom{n}{y} (1-p)^y p^{n-y}$

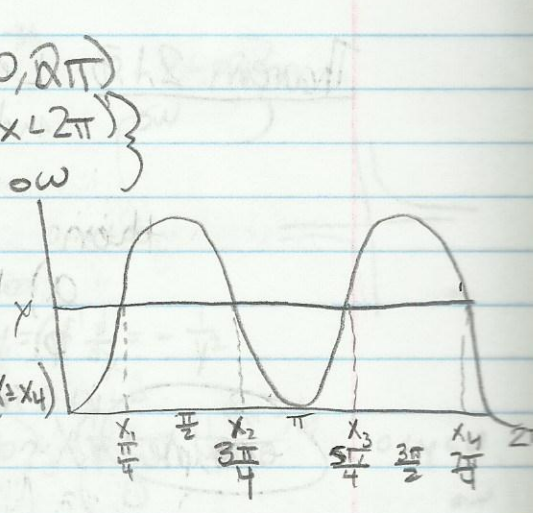
So $F_Y(y) = \int f_Y(y) = P(Y\le y) = P(g(x)\le y)$ (CDF)
$= P(\{x\in X \mid g(x)\le y\}) = \int_{x\in X \mid g(x)\le y} f_x(x)$

Example 2: Uniform: Suppose X is uniform$(0, 2\pi)$

ie $f_X(x) = \begin{cases} 1/2\pi & (0 < x < 2\pi) \\ 0 & \text{ow} \end{cases}$

- Consider $Y = \sin^2(x)$. Then

$P(Y \le y) = P(X \le x_1) + P(x_2 \le X \le x_3) + P(X \ge x_4)$



- We can say by symmetry $P(X \le x_1) = P(X \ge x_4)$

and $P(x_2 \le X \le x_3) = 2P(x_2 \le X \le \pi)$

$\therefore P(Y \le y) = 2P(X \le x_1) + 2P(x_2 \le X < \pi)$

note x_1, x_2 are $x \ni \sin^2(x) = y$ $\quad 0 < x < \pi \subseteq 0 < x < 2\pi$

Transformations w/ monotone $g(x)$:

recall:

- monotone increasing: $u > v \rightarrow g(u) > g(v)$
- monotone decreasing: $u < v \rightarrow g(u) > g(v)$

- If $g(x)$ is monotone then the mapping $x \rightarrow y$ is one to one

- monotone increasing $\rightarrow \{x \in \mathcal{X} \mid g(x) \le y\} = \{x \in \mathcal{X} \mid g^{-1}(g(x)) \le g^{-1}(y)\}$
$= \{x \in \mathcal{X} \mid x \le g^{-1}(y)\}$
- monotone decreasing $\rightarrow \{x \in \mathcal{X} \mid g(x) \le y\} = \{x \in \mathcal{X} \mid g^{-1}(g(x)) \ge g^{-1}(y)\}$
$= \{x \in \mathcal{X} \mid x \ge g^{-1}(y)\}$

monotone inc. $F_Y(y) = \int_{x \in \mathcal{X} \mid x \le g^{-1}(y)} f_X(x)\,dx = \int_{-\infty}^{g^{-1}(y)} f(x)\,dx = F_X(g^{-1}(y))$ $\quad$ b/c $\lim_{n \to -\infty} F_X(n) = 0$

monotone dec. $F_Y(y) = \int_{g^{-1}(y)}^{\infty} f_X(x)\,dx = 1 - F_X(g^{-1}(y))$

Theorem 2.1.3 Let X have cdf $F_X(x)$, let $Y = g(X)$ and let $\mathcal{X}$ and $\mathcal{Y}$ be defined:

$$\mathcal{X} = \{x \mid f_X(x) > 0\} \qquad \mathcal{Y} = \{y \mid y = g(x) \text{ for some } x \in \mathcal{X}\}$$

then

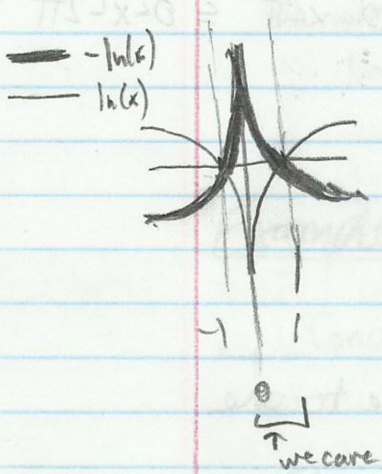
a) If g is monotone increasing on $\mathcal{X}$ $F_Y(y) = F_X(g^{-1}(y))$ for $y \in \mathcal{Y}$

b) If g is monotone increasing on $\mathcal{X}$ and X is a continuous rv $F_Y(y) = 1 - F_X(g^{-1}(y))$ for $y \in \mathcal{Y}$

example: $X \sim f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{ow} \end{cases}$ ie Uniform 0,1

$Y = g(X) = -\log(X)$

$-\log(x)\frac{d}{dx} = \frac{-1}{x} < 0 \quad \forall x$ so $g(x)$ is monotone decreasing on $(0,1)$



the support of X is $0 < x < 1$ so support of Y is $0 < y < \infty$

$$g^{-1}(y) = e^{-(y)}$$

$$F_Y(y) = 1 - F_X(g^{-1}(y)) = 1 - F_X(e^{-y}) = 1 - e^{-y} \quad \text{(b/c 2.1.3 monotone decreasing)}$$

Theorem 2.1.5. Let X have a pdf $f_X(x)$ and let $Y = g(X)$ where g is a monotone function. Let $\mathcal{X}$ and $\mathcal{Y}$ be defined

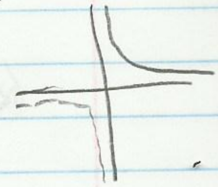
$$\mathcal{X} = \{x \mid f_X(x) > 0\} \qquad \mathcal{Y} = \{y \mid y = g(x) \text{ for some } x \in \mathcal{X}\}$$

Suppose that $f_X(x)$ is continuous on $\mathcal{X}$ and that $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$. Then the pdf of Y is given by

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| & y \in \mathcal{Y} \\ 0 & \text{ow} \end{cases}$$

Example: Let $f_X(x) = \begin{cases} \frac{1}{(n-1)!\beta^n} x^{n-1} e^{-x/\beta} & 0<x<\infty \\ 0 & ow \end{cases}$

Suppose $Y = g(X) = 1/X$



i) $g(x)$ is monotone decreasing $0<x<\infty$

ii) f_X is continuous on $0<x<\infty$

iii) $g^{-1}(y) = 1/y \quad g^{-1}(y)\, d/dy = \frac{1}{y}\frac{d}{dy} = y^{-1}\frac{d}{dy} = -\frac{1}{y^2}$

Since $y = g(x) = \frac{1}{x} \quad 0<x<\infty \rightarrow 0<y<\infty$

$$\therefore f_Y(y) = \begin{cases} \frac{1}{(n-1)!\beta^n}\left(\frac{1}{y}\right)^{n-1} e^{-(1/y)/\beta}\left|1/y^2\right| & 0<y<\infty \\ 0 & ow \end{cases} = \begin{cases} \frac{1}{\beta^n (n-1)!}\left(\frac{1}{y}\right)^{n+1} e^{-1/\beta y} & 0<y<\infty \\ 0 & ow \end{cases}$$

Example 2: X is a continuous RV. For $y>0$ the cdf of $Y=X^2$ is:

$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y})$

* B/c x is continuous we can write $= P(-\sqrt{y} < X \le \sqrt{y})$

$= P(X \le \sqrt{y}) - P(X < -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$

So

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$
$$= \frac{d}{dy}\left[F_X(\sqrt{y}) - F_X(-\sqrt{y})\right]$$
$$= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}) \quad \text{chain rule}$$
$$= \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right)$$

Theorem 2.1.8 • Let X have pdf $f_X(x)$ let $Y=g(X)$ and define the sample space $\mathcal{X}$: $\mathcal{X} = \{x \mid f_X(x) > 0\}$

• Suppose $\exists$ a partition $A_0, A_1, \dots A_k \ni P(X \in A_0) = 0$ and $f_X(x)$ is continuous on each A_i. Further suppose $\exists$ functions $g_1, g_2, g_3, \dots g_k$ defined on $A_1, A_2, A_3, \dots A_k \ni$

i) $g(x) = g_i(x)$ for $x \in A_i$

ii) $g_i(x)$ is monotone on A_i

iii) the set $\mathcal{Y} = \{y \mid y = g_i(x)$ for some $x \in A_i\}$ (the range) is the same for all A_i

iv) $g_i^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$ $\forall i$

Then $f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X(g_i^{-1}(y)) \left|g_i^{-1}(y)\frac{d}{dy}\right| & y \in \mathcal{Y} \\ 0 & ow \end{cases}$

Example: $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad -\infty < x < \infty$

$Y = g(x) = x^2$

Note: $g(x)$ is monotone decreasing $(-\infty, 0)$

increasing $(0, \infty)$

So $A_0 = \{0\}$

$A_1 = (-\infty, 0) \quad g(x) = x^2 \quad g_1^{-1}(y) = -\sqrt{y}$

$A_2 = (0, \infty) \quad g(x) = x^2 \quad g_1^{-1}(y) = \sqrt{y}$

$\therefore f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left| \frac{1}{2\sqrt{y}} \right| = \frac{1}{2\pi} \frac{1}{\sqrt{y}} e^{-y/2}$

ON $0 < y < \infty$ range of x^2 on $-\infty < x < \infty$

Theorem 2.1.10 Let X have continuous cdf $F_X(x)$ and define the random variable Y as $Y = F_X(x)$ then Y is uniformly distributed on $(0,1)$

$\boxed{2.2}$

Def 2.2.1 $E[g(x)]$ = the expected value or mean of a random variable $g(x)$

$$E[g(x)] = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x)\,dx & \text{for continuous} \\ \sum_{x \in X} g(x) f_X(x) = \sum g(x) P(X=x) & \text{for discrete} \end{cases}$$

Note: If $|E[g(x)]| = \infty$ we say $E[g(x)]$ dne

Example $F_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \quad 0 \le x < \infty \quad \lambda > 0$

$$E(x) = \int_0^\infty \frac{1}{\lambda} x e^{-x/\lambda} dx$$
$$= -xe^{-x/\lambda}\Big|_0^\infty + \int_0^\infty e^{-x/\lambda} dx \quad \text{(integration by parts)}$$
$$= \int_0^\infty e^{-x/\lambda} dx = \lambda$$

Example $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, k$

$\ni n \in \mathbb{Z}^+ \quad 0 \le p \le 1$

$$E(x) = \sum_0^n x\binom{n}{x} p^x (1-p)^{n-x} = \sum_1^n x\binom{n}{x} p^x (1-p)^{n-x} \quad \text{(b/c } x=0 \text{ term is } 0)$$
$$E(x) = \sum_1^n n\binom{n-1}{x-1} p^x (1-p)^{n-x}$$
$$= \sum_0^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y} = np$$

<u>Theorem 2.2.5</u> Let x be a RV and let a, b, c be constants. Then for any functions $g_1(x)$ and $g_2(x)$ whose expectations exist

a) $E(ag_1(x) + bg_2(x) + c) = aE(g_1(x)) + bE(g_2(x)) + c$

b) If $g_1(x) \ge 0 \ \forall x$ then $Eg_1(x) \ge 0$

c) If $g_1(x) \ge g_2(x) \ \forall x$ then $E(g_1(x)) \ge E(g_2(x))$

d) If $a \le g_1(x) \le b \ \forall x$ then $a \le Eg_1(x) \le b$

2.3

<u>Def 2.3.1</u> for each $n \in \mathbb{Z}$, the n^{th} moment of X (or $F_X(x)$), μ_n' is

$$\mu_n' = E(X^n)$$

The n^{th} central moment of X, μ_n, is

$$\mu_n = E[(X-\mu)^n] \quad \text{where } \mu = \mu_1' = E(x)$$

Mean $= E(X) = \mu_1'$

<u>Def 2.3.2</u> Variance $= E[(X - E(X))^2]$

Note if variance $= 0$ $\quad X = E(X)$ w/ probability

Example: let X have the exponential(λ) distribution defined: $f_X(x) = \frac{1}{\lambda} e^{-x/\lambda}$ $0 \le x < \infty$ $\lambda > 0$ $\Rightarrow E(X) = \lambda$

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

$$\cdots = \lambda^2$$

Theorem 2.3.4) X is a RV w/ a finite variance then for any constants and b

$$Var(aX+b) = a^2 Var(X)$$
$$= E(X^2) - (E(X))^2 \quad \text{Page 60-1}$$

Example: $X \sim$ Binomial(n,p)

ie $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$ $\quad x \in \mathbb{Z}^+$

We know $E(X) = np$

$$var(X) = E(X^2) - (np)^2$$
$$\rightarrow E(X^2) = \sum x^2 \binom{n}{x} p^x (1-p)^{n-x} = n(n-1)p^2 + np$$

thus
$$var(X) = n(n-1)p^2 + np - n^2p^2$$
$$= n^2p^2 - p^2n + np - n^2p^2$$
$$= -p^2n + np$$
$$= pn(1-p)$$

Definition 2.3.6 • Let X be a random variable with cdf F_X. The moment generating function of X (or F_X) is

$$M_X(t) = E(e^{tX})$$

• provided that the expectation exists for t in some neighborhood of 0. That is $\exists\, h > 0 \ni \forall t$ in $-h < t < h$ $E(e^{tX})$ exists. If not the MGF DNE!

$\int_{-\infty}^{\infty} e^{tx} f_X(x) dx$ Continuous

$\sum_x e^{tx} P(X=x)$ discrete

the n^{th} derivative

Theorem 2.3.7) If X has mgf $M_X(t)$ then $E(X^n) = M_X^{(n)}(0)$

$\ni M_x^n(0) = \frac{d^n}{dt^n} M_x(t)\big|_0$

✗ Example $f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot x^{\alpha-1} \cdot e^{-x/\beta}$, $0 < x < \infty$ $\alpha > 0$ $\beta > 0$

$\ni \Gamma(\alpha)$ is the gamma function

$$M_X(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty e^{tx} x^{\alpha-1} e^{-x/\beta} dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta + tx} = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x(\frac{1}{\beta} - t)}$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/(\frac{\beta}{1-\beta t})}$$ ← we make it look like a kernal (gamma)

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot \left(\Gamma(\alpha) \left(\frac{\beta}{1-\beta t}\right)^\alpha\right)$$ ← defined by gamma

$$= \left(\frac{1}{1-\beta t}\right)^\alpha \quad \text{if } t < 1/\beta$$

So $E(X) = M_x(t)\frac{d}{dt}\big|_0$

$:= (1-\beta t)^{-\alpha} \frac{d}{dt} = \alpha\beta(1-\beta t)^{-\alpha+1}\big|_0 = \alpha\beta$

✗ The set of characterizing moments is not enough to determine a distribution uniquely b/c there may be two distinct RV having the same moments ie can have all the same moments but be 2 different pdfs

Theorem 2.3.11 : Let $F_X(x)$ and $F_Y(y)$ be two cdfs all of whose moments exist

a) If X and Y have bounded support then $F_X(u) = F_Y(u)$ $\forall u$ iff $E(X^r) = E(Y^r)$ $\forall r \in \mathbb{Z}^+$

b) If the moment generating functions exist and $M_X(t) = M_Y(t)$ $\forall t$ in some neighborhood of 0 then $F_X(u) = F_Y(u)$ $\forall u$

Theorem 2.3.12 Suppose $\{X_i, i = 1, 2, \ldots\}$ is a sequence of random variables each with mgf $M_X(t)$. Further, suppose that

$\lim_{i\to\infty} M_{X_i}(t) = M_X(t)$ $\forall t$ in a neighborhood of 0

✗ That is, convergence, for $|t| < h$, of mgfs to an mgf implies convergence of cdfs

Examples 2.3.13 (Poisson Approximation of the binomial)

$X_n \sim b(150000, .01)$

$P(X \geq 100) \approx P(X \geq 100)$

$X \sim \text{Poisson}(150)$

$X_n \sim b(n,p)$

$B(n,p) \sim \text{Poisson}(\lambda)$ when n is large and np is small

$P(X=x) = (e^{-\lambda}\lambda^x)/x! \ni x \in \mathbb{Z}^+ \ni \lambda$ is a pos constant

*Approx states $X \sim \text{binomial}(n,p)$ then $Y \sim \text{Poisson}(\lambda = np)$

$P(X=x) \approx P(Y=x)$

We know $M_X(t) = [pe^t + (1-p)]^n$

$\lambda/n \, e^t + (1 - \frac{\lambda}{n})$

$(\lambda/n e^t - \frac{\lambda^2}{n^2} e^t)^n$

$M_Y(t) = e^{\lambda(e^t - 1)} \ni p = \lambda/n$

then $M_X(t) \to M_Y(t)$ as $n \to \infty$ ($\lambda/n \to p$) WOOOOO!

? **Lemma 2.3.14** Let $a_1, a_2, \ldots$ be a sequence of numbers converging to a, that is, $\lim_{n\to\infty} a_n = a$ then

$$\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^a$$

Theorem 2.3.15 For any constants a and b, the mgf of the random variable $aX+b$ is given by

$$M_{aX+b}(t) = e^{bt} M_X(at)$$

$\boxed{2.4}$ Calc Rules we should have done earlier

Theorem 2.4.1 If $f(x,\theta)\,dx = f(b(\theta),\theta)\left(\frac{d}{d\theta}b(\theta)\right) - f(a(\theta),\theta)\left(\frac{d}{d\theta}a(\theta)\right) + \int_{a(\theta)}^{b(\theta)} \frac{d}{d\theta} f(x,\theta)$

$\boxed{\text{Note}}$ if $a(\theta)$ and $b(\theta)$ are constant

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

Theorem 2.4.2 Suppose $h(x,y)$ is continuous at y_0 for each x and $\exists$ a function $g(x) \ni$

i) $|h(x,y)| < g(x)$ for all x and y

ii) $\int_{-\infty}^{\infty} g(x)\,dx < \infty$

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

Theorem 2.4.3 • Suppose $f(x,\theta)$ is differentiable at $\theta = \theta_0$, that is

$$\lim_{\delta \to 0} \frac{f(x,\theta_0+\delta) - f(x,\theta_0)}{\delta} = \frac{\partial}{\partial\theta} f(x,\theta)\Big|_{\theta=\theta_0}$$

• exists for every x, and there exist a function $g(x,\theta_0)$ and a constant $\delta_0 > 0$ $\ni$

i) $\left|\frac{f(x,\theta_0+\delta) - f(x,\theta_0)}{\delta}\right| \le g(x,\theta_0)$ $\forall x$ and $|\delta| \le \delta_0$

ii) $\int_{-\infty}^{\infty} g(x,\theta_0)\,dx < \infty$

THEN

$$\frac{d}{d\theta}\int_{-\infty}^{\infty} f(x,\theta)\,dx\Big|_{\theta=\theta_0} = \int_{-\infty}^{\infty}\left[\frac{d}{d\theta} f(x,\theta)\Big|_{\theta=\theta_0}\right]dx$$

i) can be replaced by ↓ which is usually easier to show

$$\frac{f(x,\theta_0+\delta) - f(x,\theta_0)}{\delta} = \frac{\partial}{\partial\theta} f(x,\theta)\Big|_{\theta=\theta_0+\delta(x)}$$

$\ni$ for some number $\delta^*(x)$ where $|\delta^*(x)| < \delta_0$

$\therefore$ i) will be satisfied if we find a $g(x,\theta)$ that satisfies condition (ii)

and $\left|\frac{\partial}{\partial\theta} f(x,\theta)\Big|_{\theta=\theta'}\right| \le g(x,\theta)$ for all θ' $\ni$ $|\theta'-\theta| \le \delta_0$

Corollary 2.4.4 Suppose $f(x,\theta)$ is differentiable in θ and $\exists$ a function $g(x,\theta)$ $\ni$ $\left|\frac{\partial}{\partial\theta} f(x,\theta)\Big|_{\theta=\theta'}\right| \le g(x,\theta)$ $\forall$ θ' $\ni$ $|\theta'-\theta| \le \delta_0$ and $\int_{-\infty}^{\infty} g(x,\theta)\,dx < \infty$ Then

$$\frac{d}{d\theta}\int_{-\infty}^{\infty} f(x,\theta)\,dx\Big|_{\theta=\theta_0} = \int_{-\infty}^{\infty}\frac{d}{d\theta} f(x,\theta)\Big|_{\theta=\theta_0}\,dx$$

Example $f(x) = \begin{cases} \frac{1}{\lambda} e^{-x/\lambda} & 0 < x < \infty \\ 0 & \text{ow} \end{cases}$

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

if $n > 0$

$$\frac{d}{d\lambda} E(X^n) = \int_0^\infty \frac{\partial}{\partial \lambda} x^n \left(\frac{1}{\lambda}\right) e^{-x/\lambda} \, dX$$

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

To justify the interchange of integration and differentiation we bound the derivative of $x^n (1/\lambda) e^{-x/\lambda}$.

- Now, $\left| \frac{\partial}{\partial \lambda} \left(\frac{x^n e^{-x/\lambda}}{\lambda} \right) \right| = \frac{x^n e^{-x/\lambda}}{\lambda^2} \left| \frac{x}{\lambda} - 1 \right| \leq \frac{x^n e^{-x/\lambda}}{\lambda^2} \left(\frac{x}{\lambda} + 1 \right)$ $\boxed{\text{since } \frac{x}{\lambda} > 0}$

for some constant $0 < \delta_0 < \lambda$ take

$$g(x;\lambda) = \frac{x^n e^{-x/(\lambda+\delta_0)}}{(\lambda - \delta_0)^2} \left(\frac{x}{\lambda - \delta_0} + 1 \right)$$

- We then have

$$\left| \frac{\partial}{\partial \lambda} \left(\frac{x^n e^{-x/\lambda}}{\lambda} \right) \Big|_{\lambda = \lambda'} \right| \leq g(x,\lambda) \quad \forall \lambda' \ni |\lambda' - \lambda| \leq \delta_0$$

- $\therefore \exists$ a $g(x) \ni \int_{-\infty}^\infty g(x,\lambda) \, dx < \infty$ for $\lambda - \delta_0$

We also note that the moments of the exponential distribution are recursive:

$$E(X^{n+1}) = \lambda E(X^n) + \lambda^2 \frac{d}{d\lambda} E(X^n)$$

Note: The normal distribution is similar:

$$E(X^{n+1}) = \mu E(X^n) - \frac{d}{d\mu} E(X^n)$$

Example Let X be a discrete random variable w/ a geometric distribution.

$$P(X=x) = \theta(1-\theta)^x \quad \forall\ x = 0,1,\dots \text{ and } 0<\theta<1$$

We have $\sum_0^\infty \theta(1-\theta)^x = 1$ and provided the operation is justified

$$\frac{d}{d\theta}\sum_0^\infty \theta(1-\theta)^x = \sum_0^\infty \frac{d}{d\theta}\theta(1-\theta)^x = \sum_0^\infty\left[(1-\theta)^x - \theta x(1-\theta)^{x-1}\right]$$

by splitting the sum $= \frac{1}{\theta}\sum_0^\infty \theta(1-\theta)^x - \frac{1}{1-\theta}\sum_0^\infty x\theta(1-\theta)^x$ since $\sum_0^\infty \theta(1-\theta)^x = 1\ \forall\ 0<\theta<1$ its deriv is 0,

So: $\frac{1}{\theta}\sum_0^\infty \theta(1-\theta)^x - \frac{1}{1-\theta}\sum_0^\infty x\theta(1-\theta)^x = 0$

hence: $\frac{1}{\theta} - \frac{1}{1-\theta}E(X) = 0$

so $E(X) = \frac{1-\theta}{\theta}$

Theorem 2.4.8 Suppose $\sum_0^\infty h(\theta,x)$ converges $\forall\ \theta$ in $(a,b) \subseteq \mathbb{R}$ and

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

ii) $\sum_{x=0}^\infty \frac{\partial}{\partial\theta}h(\theta,x)$ converges uniformly on every closed & bounded sub interval of (a,b) then

$$\frac{d}{d\theta}\sum_0^\infty h(\theta,x) = \sum_0^\infty \frac{\partial}{\partial\theta}h(\theta,x)$$

Example
- say $h(\theta,x) = \theta(1-\theta)^x$ and $\frac{\partial}{\partial\theta}h(\theta,x) = (1-\theta)^x - \theta x(1-\theta)^{x-1}$
- $S_n(\theta) = \sum_0^n\left[(1-\theta)^x - \theta x(1-\theta)^{x-1}\right]$ ... pg 75

Theorem 2.4.10 Suppose the series $\sum_0^\infty h(\theta,x)$ converges uniformly on $[a,b]$ and that, $\forall\ x$, $h(\theta,x)$ is a continuous function of θ then

$$\int_a^b \sum_0^\infty h(\theta,x)\,d\theta = \sum_0^\infty \int_a^b h(\theta,x)\,d\theta$$