Moments (Sect 2.6)

Def: Assume that X is a random variable and $p>0$. We say that X has <u>p-th</u> <u>moment</u> if

$$E[|X|^p] < \infty$$

If so, $m_p = E[|X|^p]$ is called the p-th moment.

Prop: If X, Y have p-th moments, then $X+Y$ has p-th moment.

Prop: If $0 \leq p \leq q$, and X has q-th moment, then X has p-th moment.

Def: If X has p-th moment, we define the <u>centered</u> <u>p-th</u> moment to be

$$E[|X - E[X]|^p]$$

Def: The <u>variance</u> of X is the centered 2nd moment

$$Var(X) = E[(X - E(X))^2]$$

Note: $Var(X) = E[X^2] - E[X]^2$

Expectations and integral

P is a prob. measure, we write

$$\underbrace{E[X]}_{\text{compact}} = \underbrace{\int X \, dP}_{\text{flexible.}}$$

Def: A is an event, then

$$\int_A X \, dP = \int \mathbb{1}_A X \, dP \qquad \mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise.} \end{cases}$$

With expectation: $\int_A X \, dP = E[\mathbb{1}_A X]$

$(\Omega, \mathcal{F}, P)$ r.v. $X : \Omega \to \mathbb{R}$ induces $(\mathbb{R}, \mathcal{B}, \mu)$

$\Downarrow$ the distribution of X.

$E_P[X] = \int X \, dP$

$$E_\mu[g] = \int g(x) \, d\mu$$

Notation $\int g(x) \, dF(x)$

distribution

Theorem: Assume that X is a random variable and $g \geq 0$, is a Borel function, then

$$E_P[g(X)] = \int g(x) \, d\mu(x) = \int g(x) \, dF(x)$$

Sketch of proof: Let g_k be the lower approximations to g:

$$g_k(x) = \frac{j}{2^k} \quad \text{if} \quad \frac{j}{2^k} < g(x) \leq \frac{j+1}{2^{k+1}}$$

Then

$$\underbrace{E_\mu[g_k]}_{\nearrow E_\mu(g)} = \sum_j \frac{j}{2^k} \mu\{x : \tfrac{j}{2^k} < g(x) \leq \tfrac{j+1}{2^k}\}$$

$$= \sum_j \frac{j}{2^k} P\{\omega : \tfrac{j}{2^k} < g(X(\omega)) \leq \tfrac{j+1}{2^k}\}$$

$$= \underline{E_P[g(X)_k]} \to E_P[g(X)]$$

Hence $E_\mu(g) = E_P[g(X)]$

$\quad \parallel$

$\int g(x) \, d\mu(x)$.

Cor: If the distribution of X has a density $f(x)$, then

$$E_\mu[g] = \underbrace{\int g(x) \, dF(x)}_{\int g(x) \, d\mu(x)} = \int g(x) f(x) \, dx$$

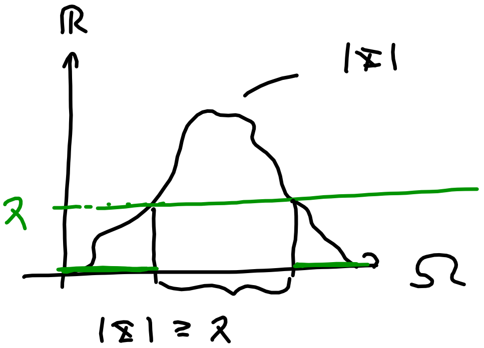
$\parallel$

$E_P[g(X)]$

Inequalities

Chebyshev's inequality: If X is a random variable and $\lambda, p > 0$, then

$$P\{|X| \geq \lambda\} \leq \frac{1}{\lambda^p} E[|X|^p]$$



Proof: $E[|X|^p] = \int |X|^p \, dP$

$$\geq \int_{\{\omega : |X| \geq \lambda\}} |\lambda|^p \, dP = |\lambda|^p P\{|X| \geq \lambda\}$$

thus $\frac{1}{|\lambda|^p} E[|X|^p] \geq P\{|X| \geq \lambda\}$

Corollary: $P[|X - E(X)| \geq \lambda] \leq \frac{1}{\lambda^2} \mathrm{Var}(X)$

Proof: Apply Chebyshev with $p = 2$ to the r.v. $X - E[X]$.

$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$

Schwarz's inequality Assume that X, Y have second moments. Then

$$E[|XY|] \leq \sqrt{E[X^2]} \sqrt{E[Y^2]}$$

Proof: For any $\lambda \in \mathbb{R}$

$$0 \leq E[(X - \lambda Y)^2] = E[X^2] - 2\lambda E[XY] + \lambda^2 E[Y^2]$$

thus

$$2\lambda E[XY] \leq E[X^2] + \lambda^2 E[Y^2] \quad \Big| : 2\lambda$$

$$E[XY] \leq \frac{1}{2\lambda} E[X^2] + \frac{\lambda}{2} E[Y^2]$$

Choose $\lambda = \frac{\sqrt{E[X^2]}}{\sqrt{E[Y^2]}}$

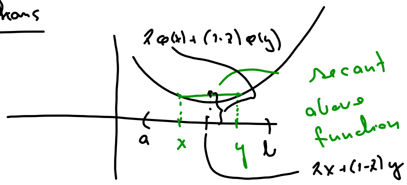
$$E[XY] \leq \frac{1}{2} \frac{\sqrt{E[Y^2]}}{\sqrt{E[X^2]}} E[X^2] + \frac{1}{2} \frac{\sqrt{E[X^2]}}{\sqrt{E[Y^2]}} E[Y^2]$$

$$= \frac{1}{2} \sqrt{E[Y^2]} \sqrt{E[X^2]} + \frac{1}{2} \sqrt{E[X^2]} \sqrt{E[Y^2]}$$

$$= \sqrt{E[X^2]} \sqrt{E[Y^2]}.$$

Convex functions

Intuition:



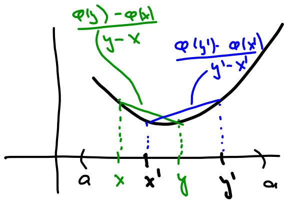
Definition: A function φ is convex on the interval (a,b) if whenever $x,y\in(a,b)$, $x<y$, then

$$\varphi(\lambda x+(1-\lambda)y)\le\lambda\varphi(x)+(1-\lambda)\varphi(y)$$

for all $\lambda\in[0,1]$.

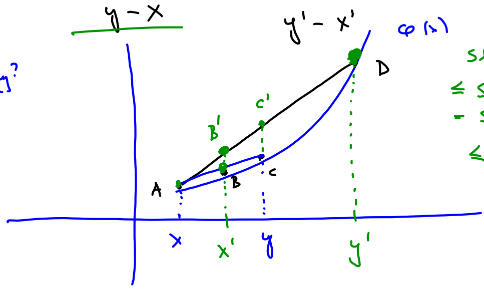
Remark: $f''(x)\ge 0$ for all $x\in(a,b)\Rightarrow f$ is convex on (a,b)

Key observation: Assume that φ is convex on (a,b). Choose points x,y and x',y' in (a,b) s.t. $x<y$, $x'<y'$ and $x<x'$ and $y<y'$



Then $\dfrac{\varphi(y)-\varphi(x)}{y-x}\le\dfrac{\varphi(y')-\varphi(x')}{y'-x'}$ $\qquad \dfrac{\varphi(y)-\varphi(x)}{x_1-x}$

Why?



Slope(AC)
$\le$ Slope(AC')
= Slope(B'D)
$\le$ Slope(BD)
$= \dfrac{\varphi(y')-\varphi(x')}{y'-x'}$

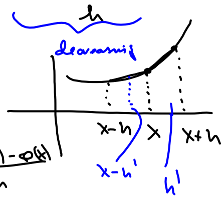
Proposition: Assume that φ is convex on (a,b). Then φ is continuous on (a,b) and it has left and right derivatives at all $x\in(a,b)$ with

$$D_-\varphi(x)\le D_+\varphi(x)$$

and both $D_-\varphi(x)$ and $D_+\varphi(x)$ are increasing functions

Proof. Let $h>0$. Then by the observation

$$\underbrace{\frac{\varphi(x)-\varphi(x-h)}{h}}_{\text{increasing}}\le\underbrace{\frac{\varphi(x+h)-\varphi(x)}{h}}_{\text{decreasing}}$$



Taking limits as $h\downarrow 0$, we get

$$D_-\varphi(x)=\lim_{h\downarrow 0}\frac{\varphi(x)-\varphi(x-h)}{h}\qquad D_+\varphi(x)=\lim_{h\downarrow 0}\frac{\varphi(x+h)-\varphi(x)}{h}$$

exist.

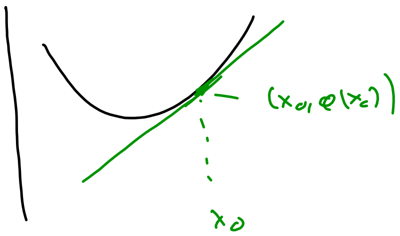
This means that φ is left- and right differentiable, hence left- and right continuous, hence continuous.

Moreover, $D_+\varphi$ is increasing since if $x<x'$

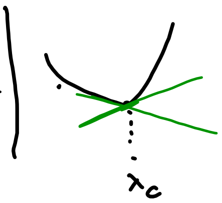
$$D_+\varphi(x)=\lim_{h\downarrow 0}\frac{\varphi(x+h)-\varphi(x)}{h}\le\lim_{h\downarrow 0}\frac{\varphi(x'+h)-\varphi(x')}{h}=D_+\varphi(x')$$

$\boxed{\text{Convention } \varphi=f}$

Supporting line in $(x_0, \varphi(x_0))$ is a line through the point such that it is always below (or on) the curve.



Proposition: Assume that φ is convex on (a,b) and that $x_0 \in (a,b)$. Then φ has a supporting line through $(x_0, \varphi(x_0))$.



Proof: Choose m such that

$$D_-\varphi(x_0) \le m \le D_+\varphi(x_0)$$

and look at the line through $(x_0, \varphi(x_0))$ with slope m:

$$y = \varphi(x_0) + m(x - x_0)$$

Assume that $x \ge x_0$: Then

$$\frac{\varphi(x) - \varphi(x_0)}{x - x_0} \ge D_+\varphi(x_0) \ge m \quad | \,(x - x_0)$$

$$\varphi(x) - \varphi(x_0) \ge m(x - x_0)$$

$$\varphi(x) \ge \underline{\varphi(x_0) + m(x - x_0)}$$

point on graph — point on line.

Assume that $x < x_0$: Then

$$\frac{\varphi(x_0) - \varphi(x)}{x_0 - x} \le D_-\varphi(x_0) \le m \quad | \, x_0 - x$$

$$\Downarrow$$

$$\varphi(x_0) - \varphi(x) \le m(x_0 - x)$$

$$\varphi(x) \ge \varphi(x_0) + m(x - x_0).$$

Jensen's inequality: Assume that φ is convex on (a,b) and X is a r.v. taking values in (a,b). Assume that X and $\varphi(X)$ are integrable. Then

$$\varphi(E[X]) \le E[\varphi(X)].$$

Proof: Choose $\underline{x_0 = E[X]}$ and let

$$y = \varphi(x_0) + m(x - x_0) \le \varphi(x) \text{ for } x \in (a,b)$$

be the supporting line.

Hence

$$\varphi(x_0) + m(\overset{(\omega)}{X} - x_0) \le \varphi(X)^{(\omega)}$$

Take expectations

$$E[\underbrace{\varphi(x_0)}_{\text{const}} + m(X - \underbrace{\overset{E[X]}{x_0}}_{\text{const}})] \le E[\varphi(X)]$$

$$\varphi(E(X)) = \varphi(x_0) + 0 \le E[\varphi(X)]$$