ecen2300

Pic 1: Intro + Reviewing what we know 29 takeaways: · correlation * causation · coverlation * causation · covariance = cov(x, z) = cov(x, z) = e(x-u_x)(z-u_z) = for							
ey takeaways: · correlation # causation · covariance = cov (x, x) = E(x-U_x)(2-U_x) = E(x) var(x) - Sample average - Substituted - Law of large numbers: Y - My increasing sample size makes Y approach population average - Central Limit Theorem (cix): Increasing sample size will make Y (sample mean) more hormally distributed. Causal effects + statistical concepts Economics suggests relationships, but never quantitative magnifudes of causal effects - in this course we want to find out Cousal effect: changing 'a' causes _ change in 'b' - we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) - so we use observational data Using observational data to estimate causal effects has difficulties: - confounding effects - simultaneous causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of Y being between values (e.g. Pr(6406) x 660)) Why use Y to estimate My? - Y is least squares estimator of My* - Y is unblased: E(Y): My - Y is consistent: Y - My - Y is consistent: Y - My - Y has smaller 6 than all other linear unblased estimators **Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuois (i.e. distance in terms of y		- INTOAN	MOTARY	I ECONU	OME TOU	C CLIMA	AADV
** takeaways: ** correlation ** causation ** correlation = corr(x, z) ** covariance = corr(x, z) ** covariance = corr(x, z) ** covariance = corr(x, z) ** correlation = corr(x, z) ** var(x) var(z) ** var(x) var(x) var(z) ** var(x) var		INTRUL	JUCIUN	ECUIN	OHE IMIC	,5 — Sulli	
** takeaways: ** correlation ** causation ** correlation = corr(x, z) ** covariance = corr(x, z) ** covariance = corr(x, z) ** covariance = corr(x, z) ** correlation = corr(x, z) ** var(x) var(z) ** var(x) var(x) var(z) ** var(x) var	OPic	1: Intro + Rev	iewing wha	t we kno	bw		
• covariance = $cov(x, x)$ var(x) var(x) = $cov(x, x)$ var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) var(x) v							
E(X-U _K)(Z-U _Z)	ey ta			160			
Law of large numbers: $\vec{y} \rightarrow \mathcal{U}_{\vec{y}}$ increasing the sample size makes \vec{y} approach population awarmage) • Central Limit Theorem (CLT): Increasing sample size will make \vec{y} (sample mean) more normally distributed. • Causal effects + statistical concepts • Conomics suggests relationships, but never quantifative magnitudes of causal effects \rightarrow : in this course we want to find out • We could use experiments to measure quantifatively • Une could use experiments to measure quantifatively • Uning observational data to estimate causal effects has difficulties: • So we use observational data • Confounding effects • Simultaneous causality • correlation \neq causation Discrete: Probability of \vec{y} being between values (e.g. $pr(\vec{y}=650)$) Continuous: Probability of \vec{y} being between values (e.g. $pr(\vec{y}=650)$) Continuous: Probability of \vec{y} being between values (e.g. $pr(\vec{y}=650)$) The substance is $\vec{y} = M_y$ • \vec{y} is unbiased: $\vec{y} = M_y$ • \vec{y} is consistent: $\vec{y} = M_y$ • \vec{y} is consistent: $\vec{y} = M_y$ • \vec{y} has smaller 6 than all other linear unbiased estimators **Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuols (i.e. distance in torms of \vec{y})		• 6	ovariance =	COV (X, 3			
Law of large numbers: $y \rightarrow My$ increasing the sample size makes $y = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{i=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{i=1}^{$)(Z-U ₃)		
Law of large numbers: $\tilde{y} \rightarrow M_y$ increasing (i.e. increasing the sample size makes \tilde{y} approach population giverage) • central limit Theorem (CLT): Increasing sample size will make \tilde{y} (sample mean) more normally distributed. Causal effects + statistical concepts Economics suggests relationships, but never quantitative magnitudes of causal effects \rightarrow : in this course we want to find out Causal effect: changing 'a' causes change in 'b' • we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) • 50 we use observational data Using observational data to estimate causal effects has difficulties: • confounding effects • simultaneous causality • correlation \neq causation Discrete: Probability of a specific value of y (e.g. $Pr(y=650)$) Continuous: Probability of y being between values (e.g. $Pr(640 \le y \le 660)$) Why use \tilde{y} to estimate $M_{\tilde{y}}$? • \tilde{y} is least squares estimator of $M_{\tilde{y}}$ * • \tilde{y} is unblased: $E(\tilde{y}) \in M_{\tilde{y}}$ • \tilde{y} is unblased: $E(\tilde{y}) \in M_{\tilde{y}}$ • \tilde{y} has smaller 6 than all other linear unblased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y			2 0000010	5220			
* Law of large numbers: Y - My summers (i.e. increasing the sample size makes Y approach population average) * Central Limit Theorem (CLT): Increasing sample size will make 9 (sample mean) more normally distributed. **Colusal effects + statistical concepts **Economics suggests relationships, but never quantitative magnitudes of causal effects - :. in this course we want to find out **Colusal effect: changing 'a' causes _ change in 'b' ** we could use experiments to measure quantitatively **But they are hardly ever feasible (expensive, time consuming etc.) ** So we use observational data **Using observational data to estimate causal effects has difficulties: ** confounding effects ** simultaneous causality ** correlation = causation ** Discrete: Probability of a specific value of y (e.g. Pr(y=650)) ** Continuous: Probability of y being between values (e.g. Pr(6405) < 660)) ** Why use \$\bar{y}\$ to estimate \$M_y\$? ** \$\bar{y}\$ is least squares estimator of \$M_y\$* ** \$\bar{y}\$ is unblased: \$E(\bar{y}) = M_y\$ ** \$\bar{y}\$ is unblased: \$E(\bar{y}) = M_y\$ ** \$\bar{y}\$ is consistent: \$\bar{y} - M_y\$ ** Yhos smaller \$\bar{y}\$ than all other linear unblased estimators ** Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuols (i.e. distance in terms of \$\bar{y}\$)						5-1 <r<13-< td=""><td>- 10</td></r<13-<>	- 10
course limit Theorem (CLT): Increasing sample size will make \$\times\$ (sample mean) more hormally distributed. Causal effects + statistical concepts Economics suggests relationships, but never quantitative magnitudes of causal effects \$\implies\$: in this course we want to find out Cousal effect: changing 'a' causes change in 'b' ' we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) ' so we use observational data Using observational data to estimate causal effects has difficulties: ' confounding effects ' simultaneous causality Continuous: Probability of a specific value of \$\text{(e.g. Pr(Y=650))}\$ Continuous: Probability of \$\text{y}\$ being between values (e.g. Pr(640 \leq \text{y} \left 660)) Why use \$\text{y}\$ to estimate \$M_{\text{y}}\$? ' \$\text{y}\$ is least squares estimator of \$M_{\text{y}}\$\$ ' \$\text{y}\$ is consistent: \$\text{y}^2 M_{\text{y}}\$\$ Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuols (i.e. distance in torms of \$\text{y}\$)	· Lo	w of large num	bers: $\bar{y} \stackrel{\rho \leftarrow}{\rightarrow}$	AL. increasi	ng size (i.E. imer	easing the same	ole size makes Ý
• central limit theorem (CLT): Increasing sample size will make \$\tilde{y}\$ (sample mean) more normally distributed. Causal effects + statistical concepts Economics suggests relationships, but never quantifative magnitudes of causal effects \(\to \): in this course we want to find out Causal effect: changing \(\tilde{a} \) causes \(\tilde{causes} \) change in \(\tilde{b} \) • we could use experiments to measure quantifatively But they are hardly ever feasible (expensive, time consuming etc.) • 30 we use observational data Using observational data to estimate causal effects has difficulties: • confounding effects • simultaneous causality • correlation \$\pm\$ causation Discrete: probability of \$\fo\$ being between values (e.g. \(\text{Pr}(\forestation) \)) Continuous: Probability of \$\fo\$ being between values (e.g. \(\text{Pr}(\forestation) \)) Why use \$\tilde{\gamma}\$ to estimate \$M_{\gamma}\$? • \$\tilde{\gamma}\$ is least squares estimator of \$M_{\gamma}\$\$ • \$\tilde{\gamma}\$ is unblased: \$E(\tilde{\gamma}) = M_{\gamma}\$\$ • \$\tilde{\gamma}\$ is consistent: \$\tilde{\gamma} = M_{\gamma}\$\$ * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$\forestation\$)					appy	oach population	average)
Causal effects + statistical concepts Economics suggests relationships, but never quantitative magnitudes of causal effects in this course we want to find out Causal effect: changing 'a' causes change in 'b' we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) so we use observational data Using observational data to estimate causal effects has difficulties: confounding effects simultaneous causality covrelation = causality covrelation = causality belong between values (e.g. Pr(Y=650)) Continuous: Probability of y belong between values (e.g. Pr(640 < Y < 660)) Why use \$\bar{Y}\$ to estimate \$M_Y\$? \$\bar{Y}\$ is least squares estimator of \$M_Y\$* \$\bar{Y}\$ is unbiased: \$E(\bar{Y}) = M_Y\$ \$\bar{Y}\$ is consistent: \$\bar{Y} = M_Y\$ \$\bar{Y}\$ is consistent: \$\bar{Y} = M_Y\$ \$\bar{Y}\$ has smaller 6 than all other linear umbiased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$Y\$)	• Ce	ntral Limit Theore	em (CLT): In	creasing s			
Economics suggests relationships, but never quantitative magnitudes of causal effects : in this course we want to find out Coursel effect: changing 'a' causes change in 'b' we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) 50 we use observational data Using observational data to estimate causal effects has difficulties: confounding effects simultaneous causality covrelation & causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\times\$ to estimate \$M_y\$? \$\times\$ is least squares estimator of \$M_y\$\$ \$\times\$ is consistent: \$\times\$ \times \$M_y\$\$ \$\times\$ is consistent: \$\times\$ \times\$ \$M_y\$\$ \$\times\$ is consistent; \$\times\$ \times\$ \$M_y\$\$ \$\times\$ is consistent; \$\times\$ \times\$ \$M_y\$\$ \$\times\$ is consistent; \$\times\$ \$M_y\$\$ \$\times\$ is consistent; \$\times\$ \times\$ \$					T I		•
Economics suggests relationships, but never quantitative magnitudes of causal effects : in this course we want to find out Cousal effect: changing 'a' causes change in 'b' • we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) • 50 we use observational data Using observational data to estimate causal effects has difficulties: • confounding effects • simultaneous causality • correlation \$\pm\$ causation Discrete: Probability of a specific value of \$\mathcal{Y}\$ (e.g. Pr(\$\mathcal{Y}\$=650)) continuous: Probability of \$\mathcal{Y}\$ being between values (e.g. Pr(\$640\$)\$=660)) Why use \$\mathcal{Y}\$ to estimate \$M_{\mathcal{Y}}\$? • \$\mathcal{Y}\$ is least squares estimator of \$M_{\mathcal{Y}}\$\$ • \$\mathcal{Y}\$ is unblased: \$E(\$\mathcal{Y}\$) = M_{\mathcal{Y}}\$\$ • \$\mathcal{Y}\$ is consistent: \$\mathcal{Y}\$ = M_{\mathcal{Y}}\$\$ • \$\mathcal{Y}\$ is consistent: \$\mathcal{Y}\$ = M_{\mathcal{Y}}\$\$ • \$\mathcal{Y}\$ has smaller \$\mathcal{S}\$ than all other linear unblased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$\mathcal{Y}\$)							
cousal effect: changing 'a' causes change in 'b' ' we could use experiments to measure quahtitatively But they are hardly ever feasible (expensive, time consuming etc.) ' so we use observational data Using observational data to estimate causal effects has difficulties: ' confounding effects ' simultaneous causality ' correlation & causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\tilde{y}\$ to estimate \$M_y\$? ' \$\tilde{y}\$ is least squares estimator of \$M_y\$* ' \$\tilde{y}\$ is unblased: \$E(\tilde{y}) = M_y\$ ' \$\tilde{y}\$ is consistent: \$\tilde{y}^2 = M_y\$ ' \$\tilde{y}\$ is consistent: \$\tilde{y}\$ is consistent \$\tilde{y}\$ is consistent; \$\tilde{y}\$ is consistent; \$\tilde{y}\$ is consistent; \$\tilde{y}\$ is consiste	Cau	isal effects +	statistical	concepts			
Cousal effect: changing 'a' causes change in 'b' ' we could use experiments to measure quahtitatively But they are hardly ever feasible (expensive, time consuming etc.) ' so we use observational data Using observational data to estimate causal effects has difficulties: ' confounding effects ' simultaneous causality ' correlation & causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\times\$ to estimate \$M_y\$? ' \$\times\$ is least squares estimator of \$M_y\$* ' \$\times\$ is unblased: \$E(\times) = M_y\$. ' \$\times\$ is consistent: \$\times\$ \times M_y\$.		• • • • • • • • • • • • • • • • • • • •					
Cousal effect: changing 'a' causes change in 'b' we could use experiments to measure quahitatively But they are hardly ever feasible (expensive, time consuming etc.) So we use observational data using observational data to estimate causal effects has difficulties: confounding effects simultaneous causality covrelation = causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) why use \$\bar{y}\$ to estimate \$M_y\$? \$\bar{y}\$ is least squares estimator of \$M_y\$\$ \$\bar{y}\$ is unbiased: \$E(\bar{y}) = M_y\$ \$\bar{y}\$ is consistent: \$\bar{y}\$ = M_y\$ \$\bar{y}\$ is consistent: \$\bar{y}\$ = M_y\$ \$\bar{y}\$ has smaller \$\bar{y}\$ than all other linear umbiased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y							udes of causal
we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) So we use observational data Using observational data to estimate causal effects has difficulties: Confounding effects simultaneous causality coverelation == causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(6405y5660)) Why use \$\bar{y}\$ to estimate \$M_y\$? \$\bar{y}\$ is least squares estimator of \$M_y\$* \$\bar{y}\$ is unbiased: \$E(\bar{y}) = M_y\$ \$\bar{y}\$ is consistent: \$\bar{y}^2 = M_y\$ \$\bar{y}\$ has smaller 6 than all other linear unbiased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y		rects -	in this c	ourse me	Want to th	na out	
• we could use experiments to measure quantitatively But they are hardly ever feasible (expensive, time consuming etc.) • So we use observational data Using observational data to estimate causal effects has difficulties: • Confounding effects • simultaneous causality • coverelation = causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(6405y5660)) Why use \$\bar{y}\$ to estimate \$M_y\$? • \$\bar{y}\$ is least squares estimator of \$M_y\$* • \$\bar{y}\$ is unbiased: \$E(\bar{y}) = M_y\$ • \$\bar{y}\$ is consistent: \$\bar{y}^{\top} M_y\$ • \$\bar{y}\$ is consistent: \$\bar{y}^{\top} M_y\$ • \$\bar{y}\$ has smaller 6 than all other linear unbiased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y	Colus	al effect: cha	naina 'a'	Causes	change	in 'h'	
Sut they are hardly ever feasible (expensive, time consuming etc.) So we use observational data Using observational data to estimate causal effects has difficulties: Confounding effects Simultaneous causality Covrelation & causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 & y & 660)) Why use \$\bar{y}\$ to estimate \$M_y\$? Y is least squares estimator of \$M_y\$* Y is unbiased: \$E(\bar{y}) = M_y\$ Y has smaller \$ than all other linear unbiased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y	•	we could use	experime	nts to m	easure aug	ntitativelu	
Using observational data to estimate causal effects has difficulties: • Confounding effects • Simultaneous causality • correlation = causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\bar{y}\$ to estimate \$M_y\$? • \$\bar{y}\$ is least squares estimator of \$M_y\$\$ • \$\bar{y}\$ is unblased: \$E(\bar{y}) = M_y\$ • \$\bar{y}\$ is consistent: \$\bar{y} = M_y\$ • \$\bar{y}\$ has smaller 6 than all other linear unblased estimators # Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y							nsuming etc.)
Using observational data to estimate causal effects has difficulties: • Confounding effects • simultaneous causality • correlation \$\pm\$ causation Discrete: Probability of a specific value of \$\forall (e.g. \text{Pr(Y=650)})\$ Continuous: Probability of \$\forall \text{being between values (e.g. Pr(640 \$\forall \text{\$<} \forall 660))} Why use \$\$\text{\$\te	•						J
• Confounding effects • simultaneous causality • correlation & causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\bar{y}\$ to estimate \$M_y\$? • \$\bar{y}\$ is least squares estimator of \$M_y\$* • \$\bar{y}\$ is unbiased: \$E(\bar{y}) = M_y\$ • \$\bar{y}\$ is consistent: \$\bar{y}^{-1} M_y\$ • \$\bar{y}\$ has smaller \$6\$ than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$y\$)							
• simultaneous causality we will unpack these later • correlation = causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 \leq y \leq 660)) Why use \$\tilde{y}\$ to estimate \$M_y\$? • \$\tilde{y}\$ is least squares estimator of \$M_y\$* • \$\tilde{y}\$ is unbiased: \$E(\tilde{y}) = M_y\$ • \$\tilde{y}\$ is consistent: \$\tilde{y} = M_y\$ • \$\tilde{y}\$ is consistent: \$\tilde{y} = M_y\$ • \$\tilde{y}\$ is consistent: \$\tilde{y} = M_y\$ • \$\tilde{y}\$ has smaller \$\tilde{x}\$ than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$y\$	Usia	ng observation	al data to	estimate	causal ef	fects has diffi	culties:
• correlation = causation Discrete: Probability of a specific value of y (e.g. Pr(y=650)) continuous: Probability of y being between values (e.g. Pr(640 < y < 660)) Why use \$\tilde{y}\$ to estimate \$M_y\$? • \$\tilde{y}\$ is least squares estimator of \$M_y\$* • \$\tilde{y}\$ is unbiased: \$E(\tilde{y}) = M_y\$ • \$\tilde{y}\$ is consistent: \$\tilde{y} = M_y\$ • \$\tilde{y}\$ has smaller \$6\$ than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$y\$				7	200		
Discrete: Probability of a specific value of y (e.g. Pr(y=650)) Continuous: Probability of y being between values (e.g. Pr(640 & y < 660)) Why use \$\bar{y}\$ to estimate \$M_y\$? • \$\bar{y}\$ is least squares estimator of \$M_y\$* • \$\bar{y}\$ is unbiased: \$E(\bar{y}) = M_y\$ • \$\bar{y}\$ is consistent: \$\bar{y}^2 - M_y\$ • \$\bar{y}\$ has smaller \$6\$ than all other linear unbiased estimators ** Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$y\$					vill unpack	these later	
continuous: Probability of Y being between values (e.g. Pr(640 < Y < 660)) Why use \$\tilde{Y}\$ to estimate \$M_{\tilde{Y}}\$? • \$\tilde{y}\$ is least squares estimator of \$M_{\tilde{Y}}\$\$ • \$\tilde{y}\$ is unbiased: \$E(\tilde{y}) = M_{\tilde{Y}}\$\$ • \$\tilde{y}\$ is consistent: \$\tilde{y}^{\tilde{Y}} = M_{\tilde{Y}}\$\$ • \$\tilde{y}\$ has smaller \$6\$ than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$Y\$)	•	correlation a	e causation				
continuous: Probability of Y being between values (e.g. Pr(640 < Y < 660)) Why use \$\bar{Y}\$ to estimate \$M_Y\$? • \$\bar{Y}\$ is least squares estimator of \$M_Y\$* • \$\bar{Y}\$ is unbiased: \$E(\bar{Y}) = M_Y\$ • \$\bar{Y}\$ is consistent: \$\bar{Y}^{\bar{Y}} \in M_Y\$ • \$\bar{Y}\$ has smaller \$6\$ than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of \$Y\$	••••				1	1 - 9-13	
Why use \bar{y} to estimate M_{γ} ? • \bar{y} is least squares estimator of M_{γ} * • \bar{y} is unbiased: $E(\bar{y}) = M_{\gamma}$ • \bar{y} is consistent: $\bar{y} - M_{\gamma}$ • \bar{y} has smaller 6 than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y	Disci	rere: Propobili	ty of a specific	being he	uue or y	(C.g. PF(7 = 6;	00))
 ÿ is least squares estimator of My* ÿ is unbiased: E(ÿ) = My ÿ is consistent: ÿ → My ÿ has smaller 6 than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y 	CONT	INGGUS. Probab	ility or y	being bes	ween value	5 (F.G. P1 (BH)	02 12 9601)
 ÿ is least squares estimator of My* ÿ is unbiased: E(ÿ) = My ÿ is consistent: ÿ → My ÿ has smaller 6 than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y 	Whe	use y to e	stimate M	.?			
 ÿ is unbiased: E(ÿ) = My ÿ is consistent: ÿ → My ÿ has smaller & than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y 	7	y is least sa	uares estim	lator of A	u.,*		
• \bar{y} is consistent: $\bar{y} \xrightarrow{\rho} M_{\gamma}$ • \bar{y} has smaller 6 than all other linear unbiased estimators ★ Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y	10	ÿ is unblase	d : E(ÿ) =)	u,			
• Ÿ has smaller 6 than all other linear unbiased estimators * Least squares estimation creates a linear regression to fit a set of data point such that the sum of the squared residuals (i.e. distance in terms of y							
such that the sum of the squared residuals (i.e. distance in terms of y					linear um	iased estimat	ors
such that the sum of the squared residuals (i.e. distance in terms of y							
							n terms of Y

```
Estimation, Hypothesis tests & Confidence Intervals
Statistical inference: using observational data to draw conclusions about the population.
  the first step to obtaining inferences about the population is estimation
Estimation: Estimators include values such as \( \) (sample mean), \( \), (15t observation),
             median, mode etc.
             4 We usually use Y
                                                  # using OLS to estimate slope
Hypothesis Testing: · null hypothesis (Ho) always includes equals (=, <,>)
                    · alternative hypothesis (H,) is an inequality (#, >, <)
Steps: O HYPOTHESIS
                                                              Ho: E(Y) >, My
                                            Ho: E(Y) & My
                        H .: E(Y) = My
                         H .: E (y) ≠ My
                                            H : E(Y) > My
                                                              H : E (Y) < My
                                             1 sided > 3
                                                              1 sided 45
                           2 sided 5
  2 DECISION RULE
                       Reject Ho if ....
                                        p < 06
                                                 (p-value method)
                                        test statistic > | critical value (t-test, 2 test)
                         t crit for a
                        t calc
                                         * | sided > *
                                         (in this case, t calc < terit : don't reject Ha)
          0
  3 TEST STATISTIC
                        test statistic = estimate - hypothesized value
                       (e.g. & cole)
                                           Std. error of estimate
      (calculate)
  (4) DECISION
                   Either reject or fail to reject Ho
                    4 Find critical value from table of critical values
  S CONCLUSION At the
                                level of significance (LOS) there is [sufficient/
                     insufficient] evidence to conclude that [statement about H,]
 Confidence Interval: A 95% confidence interval for My is an interval that
                       contains the true value of My in 95% of repeated samples.
                       Ci.e... set of values of My Not rejected by a hypothesis
                              test with a ST. LOS.
                           M: 2 ± tcrit (5/√n) p: p ± Zcrit (8p) → 8p = p(1-p)
  M: 2 + 2crit (8/m)
"based on the sample, [popn. parameter] is estimated with [LOC] 1. confidence
   to be between ___ and
Topic 2: Linear Regression with one regressor
 Key Takeaways: Y; = Bo + B, X; + u;
                                        · Bo = intercept
                                        · B, = slope
   * 'hat' (^) shows it is an estimator
                                        · Ui = regression error (omitted factors)
 R2: fraction of variance of Y explained by X (percentage, unitless)
```

```
Entity Demand OLS regression:
    Yit = B, Xit + Wit
                               where \tilde{x} denotes x minus avg. value of x (\bar{x})
                               i.e. y-y= B. (x-x) + (u-u)
LSA #1: E(U: Xiv .... Xir . &: ) = 0
Uis has mean zero, given the entity fixed effect (a) and entire history of X's

    extension of previous LSB #1

  · no omitted lagged effects
  · no feedback from u to future X
LSA # 2: (Xi, ..., Xi, ui, ..., ui, ) are j.i.d
  · satisfied if entities are randomly sampled from population using SRS
     does NOT require observations to be i.i.d. (random) for the same entity over time
     -that would be unrealistic!
     (e.g. high beer tax this year means probably high next year too -correlation)
    + cov (ut, ut+1) is often plausibly non-zero
Topic 9: Regression with a binary dependant variable
                                         -> Y=binary (i.e. = 0 or 1)
                     Y= B0+ B,X+ U
Key takeaways:
  4 can still use OLS regression as long as conditional mean independence is satisfied!
 Y
                       B. = \Delta predicted probability that Y=1, for unit \Delta X
                         but issue with using linear model is that we recieve
 0.5-
                           probabilities outside the 0-1007. range (can't interpret)
 0
                         :. we use probit or logit! (0 & pr(y=1|x) &1)
                   X
  7.
                         Probit: Pr (Y=1 | X) = (Z)
                                                             where ... = 30+B,X
                            · use normal distribution 2 table to find probability
 0.5-
 0
                         Logit: Pr (Y=1/x)= 1 + Just plug Z in to the
                                                        Lequation to find probability!
  Measures of fit for probit + logit:
  • R<sup>2</sup> and R<sup>2</sup> doesn't make sense for binary y (data points arent ON regression line)
  : instead we use...
  1. Fraction correctly predicted
  fraction of Y for which predicted prob = >507. when actual Y=1 (or <50% when Y=0)
  2. Pseudo R
   improvement in value of log-likelihood, relative to having no X's
In linear probability model, predicted value of Y is interpreted as predicted probability
 that Y=1... • If LSA #1 holds, E(Y | x) = Pr(Y=1 | x) = Bo+B.X
             · OLS is valid, given LSA's hold, and yields unbiased
For linear, probit or logit models... we can still use P-values in the same way
• ALSO, if chi<sup>2</sup> is large, it indicates statistical significance
```