

What are limit theorems?

- Limit theorems specify the probabilistic behavior of n random variables as n
- Possible restrictions on RVs:
 - Independent random variables
 - Uncorrelated random variables
 - Have identical marginal CDFs/pmfs/pmfs
 - Have identical means and/or variances

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 1 of 39

The average of n RVs

- n random variables X_1, X_2, \dots, X_n have finite expectations $\mu_1, \mu_2, \dots, \mu_n$
- Let $Z = (X_1 + X_2 + \dots + X_n)/n$
- What is $E[Z]$?
- Expectation is a **linear operator**
- $E[Z] = (E[X_1] + E[X_2] + \dots + E[X_n])/n$
- Expected value of average of n RVs = **numerical average** of their expectations
- If $E[X_i] = \mu$ for all i , then $E[Z] = \mu$ also

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 2 of 39

The sample mean

- Model: An experiment is repeated n times
- X_1, X_2, \dots, X_n are the n observed values of a random variable X on the n independent trials of the experiment
- Random variable X has **finite mean** μ
- The X_i 's are said to be **independent identically distributed** (i.i.d. or iid) random variables
- $Z = (X_1 + X_2 + \dots + X_n)/n$ is called the **sample mean**

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 3 of 39

Variance of the sample mean

- i.i.d. RVs X_i with **finite mean and variance**
- Sample mean $Z = (X_1 + X_2 + \dots + X_n)/n$
- $E[Z] = E[X] = \mu$
- $\text{var}(Z) = n^{-2} \cdot \text{var}(X_1 + X_2 + \dots + X_n)$

$$= n^{-2} \cdot [\text{var}(X_1) + \text{var}(X_2) + \dots + \text{var}(X_n)]$$

$$= n^{-1} \cdot \text{var}(X)$$
- This is because the RVs are independent
- Hence, $\text{cov}(X_i, X_j) = 0$ if $i \neq j$
- Also holds if the RVs are **uncorrelated**

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 4 of 39

Variance decreases as n increases

- Z is the **average** of the n observed values of a random variable X with mean μ and variance $\text{var}(X)$
- $E[Z] = \mu$ and $\text{var}(Z) = n^{-1} \cdot \text{var}(X)$
- If we wish to **estimate** the value of μ , then the value of the **sample mean Z** is a **much better estimator** than the value of any individual observation X_i
- Application to experimental results

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 5 of 39

Confidence interval for mean

- $E[Z] = \mu$ and $\text{var}(Z) = n^{-1} \cdot \text{var}(X) = n^{-1} \cdot \sigma^2$
- Assume $\text{var}(X) = \sigma^2$ is known
- Chebyshev inequality:

$$P\{ |X - \mu| \geq a \} \leq (\sigma/a)^2$$
- $X \pm 5\sigma$ is a 96% confidence interval for μ
- $P\{ |Z - \mu| \geq a \} \leq \sigma^2 / (a^2 n) = (\sigma/a\sqrt{n})^2$
- $Z \pm 5\sigma/\sqrt{n}$ is 96% confidence interval for μ
- **Much smaller** confidence interval with sample mean than with any individual X_i

ECE 313 - Lecture 41 © 2000 Dilip V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 6 of 39

Normality increases confidence!!

- Assume the X_i are Gaussian RVs
- For any confidence level, we get a much **smaller** confidence interval
- If the X_i are Gaussian RVs, then so is Z
- $E[Z] = \mu$ and $\text{var}(Z) = n^{-1} \cdot \text{var}(X) = n^{-1} \cdot \sigma^2$
- $(1.96) \sigma / n$ is a 95% confidence interval for μ
- $Z \pm 1.96 \sigma / n$ is a 95% confidence interval for μ
- $Z \pm 5 \sigma / n$ is a 99.99997...% confidence interval for μ

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 7 of 39

The sample variance

- Sample mean** $Z = (X_1 + X_2 + \dots + X_n)/n$ is a RV that is a function of the X_i 's
- i-th deviation** is $X_i - Z$. $\text{cov}(X_i - Z, Z) = 0$
- Sample variance** $S^2 = (n-1)^{-1} \cdot \sum (X_i - Z)^2$
- S^2 is also a RV
- Distinguish between the **sample variance** S^2 and the **variance of the sample mean**
- S^2 is a **random variable**: the variance of the sample mean is the **number** $\text{var}(X)/n$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 8 of 39

Expectation of the sample variance

- $S^2 = (n-1)^{-1} \cdot \sum (X_i - Z)^2$
- $E[S^2] = (n-1)^{-1} \cdot E[\sum (X_i - Z)^2]$
- But, $X_i - Z$ is a zero-mean random variable (why?) and hence we have that $E[S^2] = (n-1)^{-1} \cdot \text{var}(\sum (X_i - Z))$
- $= (n-1)^{-1} \cdot \text{var}(X_1) + \text{var}(Z) - 2 \cdot \text{cov}(X_1, Z)$
- $= (n-1)^{-1} \cdot \text{var}(X) + \text{var}(X)/n - 2 \cdot \text{cov}(X_1, Z)/n$ since $\text{cov}(X_i, n^{-1} \sum X_j) = n^{-1} \cdot \text{cov}(X_i, X_i)$
- $E[S^2] = (n-1)^{-1} \cdot ((n-1)\text{var}(X)) = \text{var}(X)$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 9 of 39

Variance of the sample variance S^2

- If $\text{var}(X)$ is unknown, it can be estimated as $S^2 = (n-1)^{-1} \cdot \sum (X_i - Z)^2$
- $E[S^2] = \text{var}(X)$, the unknown quantity
- The **variance** of the random variable S^2 is $n^{-1} \cdot \{E[(X-\mu)^4] - (n-3)(\text{var}(X))^2/(n-1)\}$ and also decreases as n increases
- Hence, we can use the Chebyshev inequality to find confidence intervals for the unknown variance $\text{var}(X)$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 10 of 39

Gaussian samples — I

- If the X_i are i.i.d. $N(\mu, \sigma^2)$ RVs, then
 - $Z = (X_1 + X_2 + \dots + X_n)/n$ is $N(\mu, \sigma^2/n)$
 - If σ^2 is known, then $Z \pm 1.96 \sigma / n$ is a 95% confidence interval for μ
 - $S^2 = (n-1)^{-1} \sum (X_i - Z)^2$ is a **gamma RV** with parameters $((n-1)/2, (n-1)\sigma^2/2)$ with mean σ^2 and variance $2\sigma^4/(n-1)$
 - Z and S^2 are **independent** RVs
 - Each **deviation** $X_i - Z$ is **independent** of the sample mean Z

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 11 of 39

Gaussian samples — II

- If the X_i are i.i.d. $N(\mu, \sigma^2)$ RVs, then
 - $S^2 = (n-1)^{-1} \sum (X_i - Z)^2$ is a **gamma RV** with mean σ^2 and variance $2\sigma^4/(n-1)$
 - $(n-1)S^2/\sigma^2 = \sum (X_i - Z)^2/\sigma^2$ is a χ^2 RV with $n-1$ degrees of freedom = $((n-1)/2, 1/2)$ RV with mean $n-1$
 - The value of the sample variance S^2 is a good estimate of σ^2
 - Confidence intervals and levels can be set using known gamma pdf

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 12 of 39

Gaussian samples — III

- If the X_i are i.i.d. $N(\mu, \sigma^2)$ RVs, then
 - $S = [(n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2]^{1/2}$
 - $(\bar{X} - \mu)/(S/\sqrt{n})$ has Student's t distribution with $n-1$ degrees of freedom
 - pdf is $C \cdot [1 + u^2/(n-1)]^{-n/2}$, $-\infty < u < \infty$ where $C = \frac{\Gamma(n/2)}{\sqrt{\pi} \Gamma(n-1)/2} ((n-1))^{-1/2}$
 - If $n > 2$, mean 0 and variance $n/(n-2)$
 - When σ^2 is unknown, confidence levels and intervals for mean μ are set using the t distribution

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Weak Law of Large Numbers — I

- Weak Law of Large Numbers:
 - If $X_1, X_2, \dots, X_n, \dots$ are i.i.d. RVs with finite mean μ , then for every $\epsilon > 0$,

$$P\{|\sum_{i=1}^n X_i/n - \mu| < \epsilon\} \rightarrow 1 \text{ as } n \rightarrow \infty$$
 - Equivalently

$$P\{|\sum_{i=1}^n X_i/n - \mu| < \epsilon\} \rightarrow 1 \text{ as } n \rightarrow \infty$$
 - Note that it is **not necessary** for the RVs to have finite variance
 - But the proof is easier if variance is finite

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Weak Law of Large Numbers — II

- Weaker Law of Large Numbers:
 - $X_1, X_2, \dots, X_n, \dots$ are i.i.d. with finite mean μ and finite variance σ^2 . For every $\epsilon > 0$,

$$P\{|\sum_{i=1}^n X_i/n - \mu| < \epsilon\} \rightarrow 1 \text{ as } n \rightarrow \infty$$
 - $E[\sum_{i=1}^n X_i/n] = \mu$
 - $\text{var}(\sum_{i=1}^n X_i/n) = \sigma^2/n$
 - Chebyshev inequality gives

$$P\{|\sum_{i=1}^n X_i/n - \mu| < \epsilon\} \geq 1 - (\sigma^2/n)/\epsilon^2$$
 which converges to 1 as $n \rightarrow \infty$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Weak Law of Large Numbers — III

- Even More Weaker Law of Large Numbers
 - $X_1, X_2, \dots, X_n, \dots$ are uncorrelated random variables with finite mean μ and finite variance σ^2 . For every $\epsilon > 0$,

$$P\{|\sum_{i=1}^n X_i/n - \mu| < \epsilon\} \rightarrow 1 \text{ as } n \rightarrow \infty$$
 - Weak law applies because the key result that $\text{var}(\sum_{i=1}^n X_i/n) = \sigma^2/n$ requires only that the RVs be uncorrelated: it is not necessary that the RVs be independent

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Weak Law of Large Numbers — IV

- Weak Law of Large Numbers cannot be used if $E[X_i]$ is not finite or is not defined
- Suppose that the X_i are i.i.d. Cauchy RVs
- What is the pdf of $(X_1 + X_2 + \dots + X_n)/n$?
- Characteristic function of X_i is $\exp(-|t|)$
- Characteristic function of $X_1 + X_2 + \dots + X_n$ is $\exp(-n|t|)$
- $(X_1 + X_2 + \dots + X_n)/n$ is also a Cauchy RV
- $P\{|\sum_{i=1}^n X_i/n| < \epsilon\}$ is same for all n !

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Strong Law of Large Numbers — I

- Strong Law of Large Numbers:
 - If $X_1, X_2, \dots, X_n, \dots$ are i.i.d. RVs with finite mean μ , then

$$P\{\lim_{n \rightarrow \infty} \sum_{i=1}^n X_i/n = \mu\} = 1$$
 - What's going on?
 - Isn't that just you said the Weak Law of Large Numbers is?
 - No: WLLN says $\lim_{n \rightarrow \infty} P\{\text{something}\} = 1$
 - SLLN says $P\{\lim_{n \rightarrow \infty} \text{something}\} = 1$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 14 of 39

Strong Law of Large Numbers — II

- Strong Law of Large Numbers:
If $X_1, X_2, \dots, X_n, \dots$ are i.i.d. RVs with finite mean μ , then
 $P\{\lim_n (X_1 + X_2 + \dots + X_n)/n = \mu\} = 1$
- Experiment will be repeated infinitely often
- We observe that the RV X took on values $x_1, x_2, \dots, x_n, \dots$ on these trials
- What can be said about $(x_1 + x_2 + \dots + x_n)/n$?
- We think this should be approximately μ

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 19 of 39

Strong Law of Large Numbers — III

- Experiment will be repeated infinitely often
- We observe that the RV X took on values $x_1, x_2, \dots, x_n, \dots$ on these trials
- What can be said about the sequence whose n-th term is $(x_1 + x_2 + \dots + x_n)/n$?
- There are three possibilities
 - Sequence converges to μ
 - or it converges to some other number
 - or it does not converge at all

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 20 of 39

Strong Law of Large Numbers — IV

- The sequence whose n-th term is $(x_1 + x_2 + \dots + x_n)/n$ either
 - converges to μ
 - or it converges to some other number
 - or it does not converge at all
- The Strong Law of Large Numbers says that $P\{(x_1 + x_2 + \dots + x_n)/n \text{ converges to } \mu\} = 1$
- $P\{(x_1 + x_2 + \dots + x_n)/n \text{ does not converge at all or converges to some other number}\} = 0$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 21 of 39

Strong Law of Large Numbers — V

- If the Strong Law of Large Numbers holds, then so does the Weak Law
- In fact, both require only that the RVs be i.i.d. with finite mean μ
- But, the Weak Law of Large Numbers might be applicable in cases when the Strong Law does not hold
- Example: Weak Law of Large Numbers still applies if the RVs are uncorrelated but not independent

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 22 of 39

What's a limit, anyway?

- A sequence $t_1, t_2, \dots, t_n, \dots$ converges to the limit L if for every choice of $\epsilon > 0$, we can find an integer N such that
 $|t_n - L| < \epsilon$ for all $n > N$
- Put another way, for each choice of ϵ , there can only be finitely many integers n for which the inequality $|t_n - L| < \epsilon$ does not hold
- In fact, there are obviously fewer than N such integers

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 23 of 39

Limit of a sequence of RVs

- Let $Z_n = (X_1 + X_2 + \dots + X_n)/n$ where the X_i are i.i.d. RVs with finite mean μ
- The value of Z_n is approximately μ
- For each fixed choice of ϵ , there will be some (possibly infinitely many) values of n for which Z_n is close to μ , and the event $|Z_n - \mu| < \epsilon$ will have occurred
- For other (possibly infinitely many) values of n , the event $|Z_n - \mu| > \epsilon$ will have occurred

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 24 of 39

What the Strong Law means

- Let $Z_n = (X_1 + X_2 + \dots + X_n)/n$ where the X_i are i.i.d. RVs with finite mean μ
- For each fixed choice of ϵ ,
 - $P\{\text{event } |Z_n - \mu| < \epsilon \text{ occurs for infinitely many choices of } n\} = 1$
 - $P\{\text{event } |Z_n - \mu| > \epsilon \text{ occurs for finitely many choices of } n\} = 0$
 - $P\{\text{event } |Z_n - \mu| > \epsilon \text{ occurs for infinitely many choices of } n\} = 0$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 25 of 39

Strong Law versus Weak Law

- The Strong Law of Large Numbers says
 - $P\{\text{event } |Z_n - \mu| > \epsilon \text{ occurs for finitely many choices of } n\} = 1$
 - $P\{\text{event } |Z_n - \mu| > \epsilon \text{ occurs for infinitely many choices of } n\} = 0$
- The Weak Law does not require that the event $|Z_n - \mu| > \epsilon$ occurs only for finitely many choices of n — the event might occur infinitely often: just so long as $P\{|Z_n - \mu| > \epsilon\} \rightarrow 0$ as $n \rightarrow \infty$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 26 of 39

Strong Law and relative frequencies

- The Strong Law of Large Numbers justifies the estimation of probabilities in terms of relative frequencies
- If the X_i are i.i.d. Bernoulli RVs with parameter p (and hence, finite mean p), then the sample mean Z_n converges to p with probability 1 as $n \rightarrow \infty$
- The observed relative frequency of an event of probability p converges to p with probability 1 as $n \rightarrow \infty$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 27 of 39

The Central Limit Theorem — I

- The Strong Law of Large Numbers says that the sample mean Z_n converges to μ with probability 1
- The CDF of Z_n converges to a unit step function with 0-1 transition at the point μ
- Suppose the RVs X_i have variance σ^2
- Then, $E[Z_n] = \mu$, $\text{var}(Z_n) = \sigma^2/n$
- $Y_n = n \cdot (Z_n - \mu) = (X_1 + X_2 + \dots + X_n - n\mu)$ is a RV with mean 0 and variance $n\sigma^2$
- What can we say about the CDF of Y_n ?

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 28 of 39

The Central Limit Theorem — II

- $Y_n = n \cdot (Z_n - \mu) = (X_1 + X_2 + \dots + X_n - n\mu)$ is a RV with mean 0 and variance $n\sigma^2$
- The Central Limit Theorem asserts for large values of n that the CDF of Y_n is well-approximated by the unit Gaussian CDF $\Phi(\cdot)$
- Formally, the Central Limit Theorem states that the CDF converges to $\Phi(\cdot)$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 29 of 39

The Central Limit Theorem — III

- Central Limit Theorem: Given i.i.d. RVs X_i with finite mean μ and finite variance σ^2 , the CDF of the RV $(X_1 + X_2 + \dots + X_n - n\mu) / (n\sigma)$ converges to the unit Gaussian CDF $\Phi(\cdot)$, that is, for each choice of real number u , $P\{(X_1 + \dots + X_n - n\mu) / (n\sigma) \leq u\} \rightarrow \Phi(u)$ as $n \rightarrow \infty$
- The proof uses the moment-generating function (or characteristic function) of the X_i and will be omitted

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 30 of 39

The Central Limit Theorem — IV

- The Central Limit Theorem **does not** claim that

$$Y_n = (X_1 + X_2 + \dots + X_n - n\mu) / \sqrt{n}$$

- is a unit Gaussian random variable
- All the Central Limit Theorem says is that $P\{Y_n \leq u\}$ is approximately $\Phi(u)$ when n is large
- Thus, we have a **computational tool** for probabilities involving the sample mean

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 31 of 39

The Central Limit Theorem — V

- In practical use of the Central Limit Theorem, we hardly ever use the RV

$$Y_n = (X_1 + X_2 + \dots + X_n - n\mu) / \sqrt{n}$$

- Instead, $X_1 + X_2 + \dots + X_n$ is treated as if its CDF is approximately that of a $N(n\mu, n^2)$ RV
- Thus, we compute $P\{X_1 + X_2 + \dots + X_n \leq u\} \approx \Phi((u - n\mu) / \sqrt{n})$ which is effectively the same computation

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 32 of 39

Use and Abuse of the CLT — I

- $P\{X_1 + X_2 + \dots + X_n \leq u\} \approx \Phi((u - n\mu) / \sqrt{n})$ gives a good approximation when $u \approx n\mu$
- More generally, $P\{a \leq X_1 + X_2 + \dots + X_n \leq b\} \approx \Phi((b - n\mu) / \sqrt{n}) - \Phi((a - n\mu) / \sqrt{n})$ is a **good approximation** if $a < n\mu < b$
- It is a very poor approximation if $a \gg n\mu$ or $b \ll n\mu$
- Slow convergence in the **tails** of the CDF

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 33 of 39

Use and Abuse of the CLT — II

- The RV $X_1 + X_2 + \dots + X_n$ **should not** be treated as a Gaussian RV with all the rights and privileges appertaining thereto
- Example: $X_1 + X_2 + \dots + X_n$ and $X_{n+1} + X_{n+2} + \dots + X_{2n}$ are independent RVs, and each has a CDF that is **approximately** Gaussian, but their **joint CDF is not jointly Gaussian CDF**
- Such false premises lead to false conclusions

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 34 of 39

How to Use the CLT

- $P\{a \leq X_1 + X_2 + \dots + X_n \leq b\} \approx \Phi((b - n\mu) / \sqrt{n}) - \Phi((a - n\mu) / \sqrt{n})$ is a **good approximation** if $a < n\mu < b$
- It is a very poor approximation if $a \gg n\mu$ or $b \ll n\mu$
- We saw this already for the case of binomial RVs with the DeMoivre-LaPlace theorem
- There are many versions of the CLT

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 35 of 39

Summary — I

- We learned about the properties of sample means and sample variances
- We learned about the pdfs of the sample mean and sample variances for Gaussian samples
- Weak Law of Large Numbers: If $X_1, X_2, \dots, X_n, \dots$ are i.i.d. RVs with **finite mean** μ , then for every $\epsilon > 0$, $P\{|(X_1 + X_2 + \dots + X_n) / n - \mu| \leq \epsilon\} \rightarrow 1$ as $n \rightarrow \infty$

ECE 313 - Lecture 41 © 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign, All Rights Reserved Slide 36 of 39

Summary — II

- Strong Law of Large Numbers: If $X_1, X_2, \dots, X_n, \dots$ are i.i.d. RVs with finite mean μ , $P\{\lim_n (X_1 + X_2 + \dots + X_n)/n = \mu\} = 1$
- If the Strong Law holds, then so does the Weak Law but the Weak Law may hold in cases where the Strong Law does not
- We attempted to understand the difference in what the two Laws are saying
- SLLN justifies estimation of probabilities in terms of relative frequencies

ECE 313 - Lecture 41

© 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign. All Rights Reserved.

Slide 37 of 39

Summary — III

- Central Limit Theorem: Given i.i.d. RVs X_i with finite mean μ and finite variance σ^2 , the CDF of the RV $(X_1 + X_2 + \dots + X_n - n\mu)/\sigma\sqrt{n}$ converges to the unit Gaussian CDF $\Phi(\bullet)$
- $P\{a \leq X_1 + X_2 + \dots + X_n \leq b\} = \Phi((b - n\mu)/\sigma\sqrt{n}) - \Phi((a - n\mu)/\sigma\sqrt{n})$ is a good approximation if $a < \mu < b$ and a bad approximation if $a \gg \mu$ or $b \ll \mu$

ECE 313 - Lecture 41

© 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign. All Rights Reserved.

Slide 38 of 39

Final Remarks

"We see that the theory of probability is at bottom only common sense reduced to calculation; it makes us appreciate with exactitude what reasonable minds feel by a sort of instinct, often without being able to account for it. ... It is remarkable that this science, which originated in the consideration of games of chance, should become the most important object of human knowledge. ... The most important questions of life are, for the most part, really only problems of probability."
 Pierre Simon, Marquis de LaPlace, *Analytical Theory of Probability*

ECE 313 - Lecture 41

© 2000 Dilly V. Sarwate, University of Illinois at Urbana-Champaign. All Rights Reserved.

Slide 39 of 39