

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Mathematics for Computer Science
MIT 6.042J/18.062J

Independent Sampling Theorem



Albert R Meyer,

May 13, 2013

sampletheorem.1

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Weak Law of Large Numbers

A_n ::= avg of n indep RV's
with mean μ

Theorem: For all $\delta > 0$

$$\lim_{n \rightarrow \infty} \Pr[|A_n - \mu| > \delta] = 0$$

Proof:



Albert R Meyer,

May 13, 2013

sampletheorem.2

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Weak Law of Large Numbers

A_n ::= avg of n indep RV's
with mean μ , var σ^2

Theorem: For all $\delta > 0$

$$\lim_{n \rightarrow \infty} \Pr[|A_n - \mu| > \delta] = 0$$

Proof:



Albert R Meyer,

May 13, 2013

sampletheorem.3

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Repeated Trials

$$\begin{aligned} E[A_n] &::= E\left[\frac{R_1 + R_2 + \dots + R_n}{n}\right] \\ &= \frac{E[R_1] + E[R_2] + \dots + E[R_n]}{n} \\ &= \frac{n\mu}{n} = \mu \end{aligned}$$



Albert R Meyer,

May 13, 2013

sampletheorem.4

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Weak Law of Large Numbers

So by Chebyshev

$$\Pr[|A_n - \mu| > \delta] \leq \frac{\text{Var}[A_n]}{\delta^2}$$

need only show

$$\text{Var}[A_n] \rightarrow 0 \text{ as } n \rightarrow \infty$$



Albert R Meyer,

May 13, 2013

sampletheorem.5

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Var[A_n]

$$\begin{aligned} \text{Var}[A_n] &= \text{Var}\left[\frac{R_1 + R_2 + \dots + R_n}{n}\right] \\ &= \frac{\text{Var}[R_1] + \text{Var}[R_2] + \dots + \text{Var}[R_n]}{n^2} \end{aligned}$$

$$\text{QED} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \rightarrow 0$$



Albert R Meyer,

May 13, 2013

sampletheorem.7

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Analysis of the Proof

proof only used that R_1, \dots, R_n have

- same mean
- same variance
- & variances add
— which follows from pairwise independence



Albert R Meyer,

May 13, 2013

sampletheorem.8

6	9	13	7
12	10	5	
3	1	4	14
15	8	11	2

Pairwise Independent Sampling

Theorem:

Let R_1, \dots, R_n be pairwise independent random vars with the same finite mean μ and variance σ^2 . Let

$A_n ::= (R_1 + R_2 + \dots + R_n) / n$. Then

$$\Pr[|A_n - \mu| > \delta] \leq \frac{1}{n} \left(\frac{\sigma}{\delta} \right)^2$$



Albert R Meyer,

May 13, 2013

sampletheorem.9

6	9	13	7
12		10	5
3	1	4	14
15	8	11	2

Pairwise Independent Sampling

The punchline:

we now know how big a sample is needed to estimate the mean of any* random variable within any* desired tolerance with any* desired probability

*variance $< \infty$, tolerance > 0 ,
probability < 1

