Re-Define the Test Coverage

by “Law Of Large Numbers”

大数定律重塑软件测试基础

Agenda

- Review Software Test Coverage
  - Definition Now
  - Real Case
- Re-define by Hoeffding Inequality
  - Confidence Interval
  - Confidence Level
  - Hoeffding Inequality
- Utilization of Hoeffding Inequality in Software Testing
  - Decision before Release
  - Test Estimation in Planning
- Future Application
- Appendix: Proven Real Project Data

Software Test Coverage Definition Now

White Box Test Coverage

$$\frac{\text{Code line/branch/decision required to cover}}{\text{Total code line/branch/decision}} \%$$

Black Box Test Coverage

$$\frac{\text{Coverage Item in testing}}{\text{Total Coverage Item}} \%$$

100% Coverage

+

100% Pass

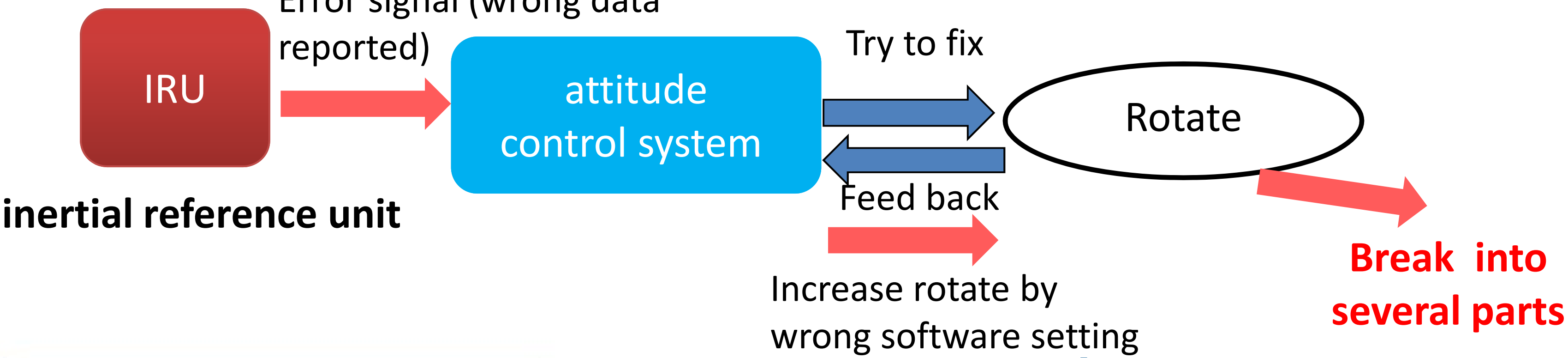
≠ ?

"Perfect" Software System

Real Case 1

The spacecraft **was launched on 17 February 2016** and **contact was lost on 26 March 2016**, due to multiple incidents with the attitude control system leading to an uncontrolled spin rate and breakup of structurally weak elements.

Hitomi X-ray astronomy satellite



Real Case 2

Software development engineers have made a mistake in the coding of computer systems, which has led to a series of problems.

In 2008, 12 "Raptor" executive from Hawaii flew to Japan's mission, when passing through the international dateline, aircraft on the global positioning system (GPS) are a failure, multiple computer system crash, restart several times also were unsuccessful. There is no way the pilots can correctly identify the location of the fighter, the altitude and speed of the aircraft.

Review:
Software Test Coverage Definition

The Base of Software Testing: Coverage->is test enough?

100% Coverage

+

100% Pass

$\neq$?

"Perfect"
Software System

100% pass in test

$\bar{X}$

In Sample

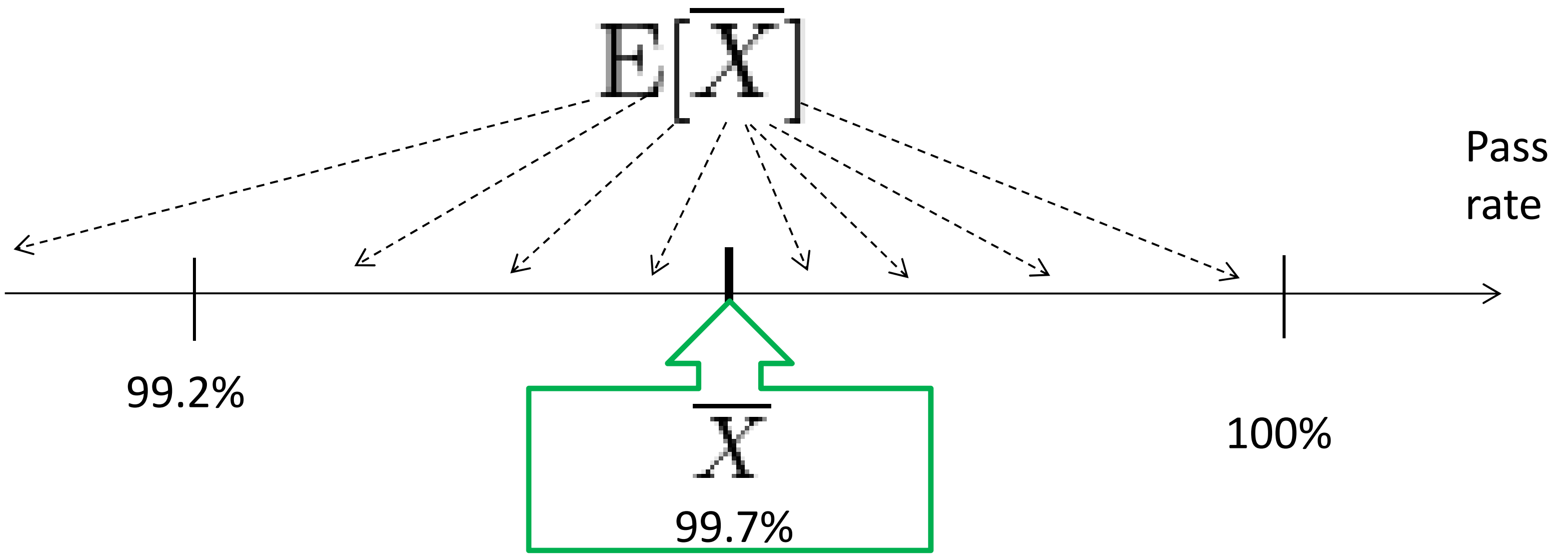
$\neq$

100% pass in market

$\mathrm{E}[\bar{X}]$

Out of Sample

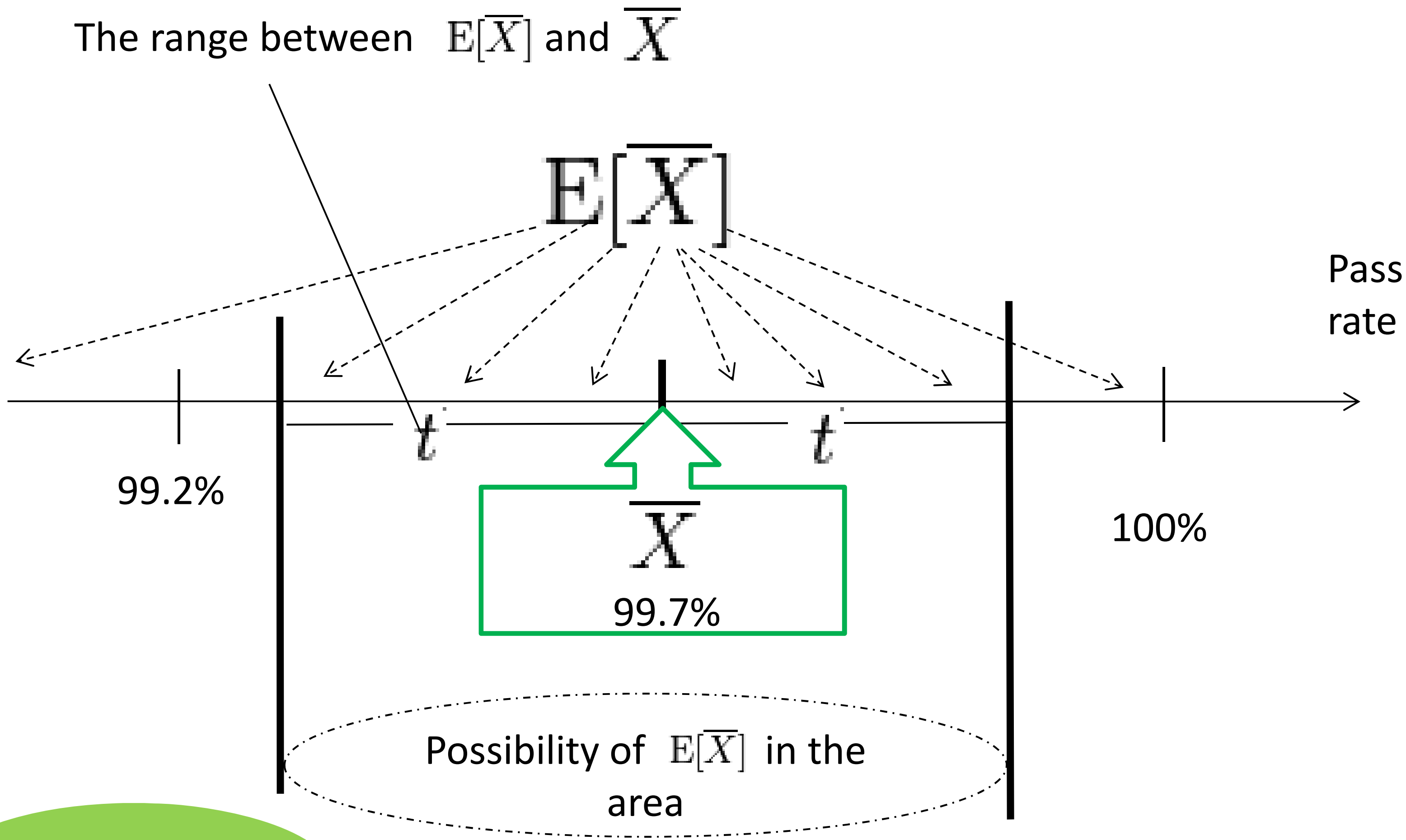
Redefine-Coverage



Coverage

How $E[\overline{X}]$ related to $\overline{X}$?

Redefine-Coverage

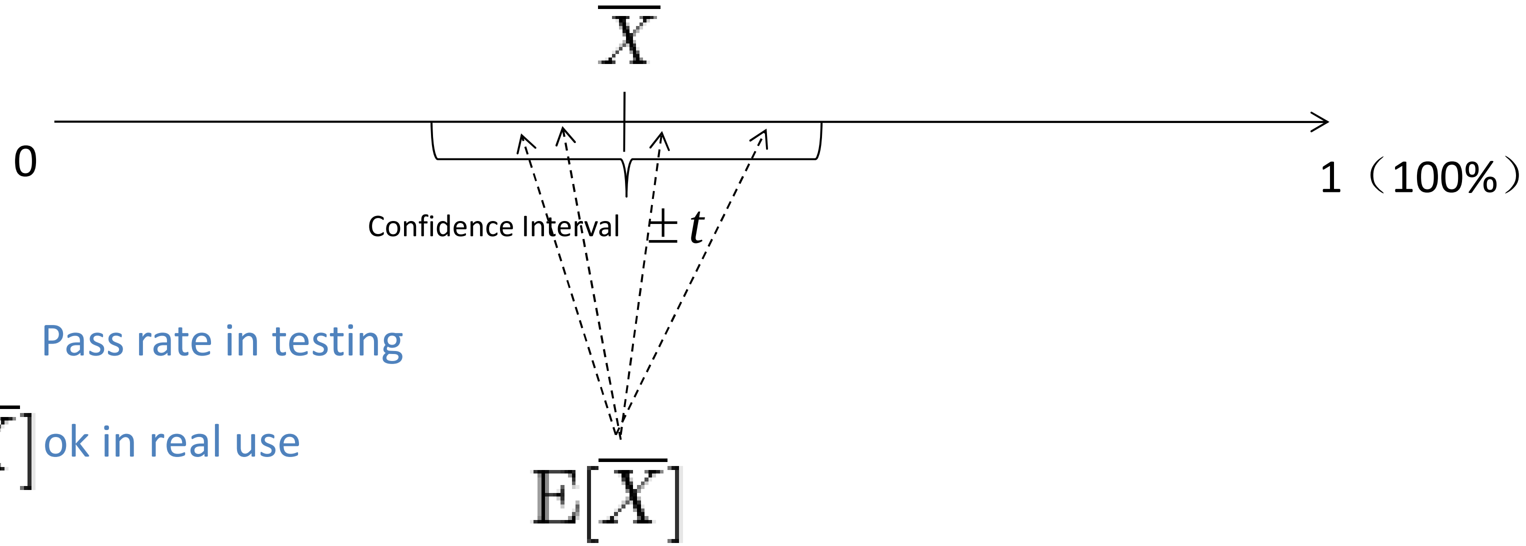


Coverage

How $E[\overline{X}]$ Close to $\overline{X}$?

Redefine-Coverage
Confidence Interval

- t --Confidence Interval



$\overline{X}$ Pass rate in testing

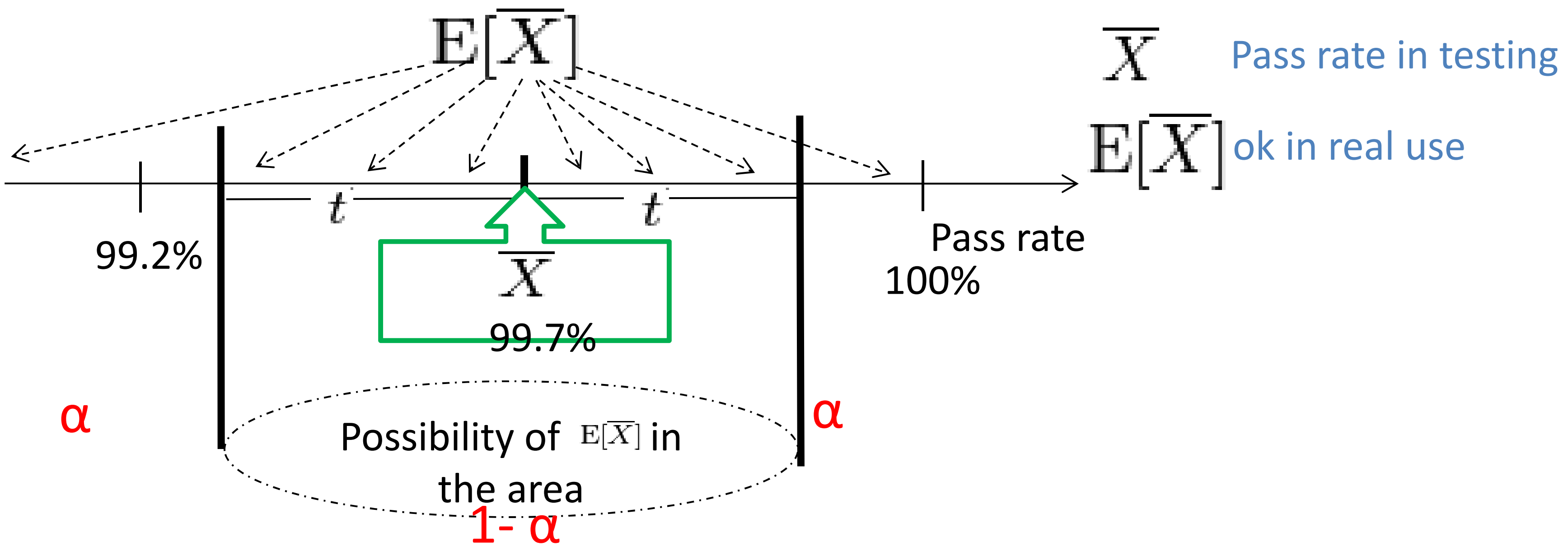
$\mathrm{E}[\overline{X}]$ ok in real use

Coverage

How $\mathrm{E}[\overline{X}]$ **Close** to $\overline{X}$?

Small t

Redefine-Coverage
Confidence Level



- "1- α" -- Confidence Level

Coverage

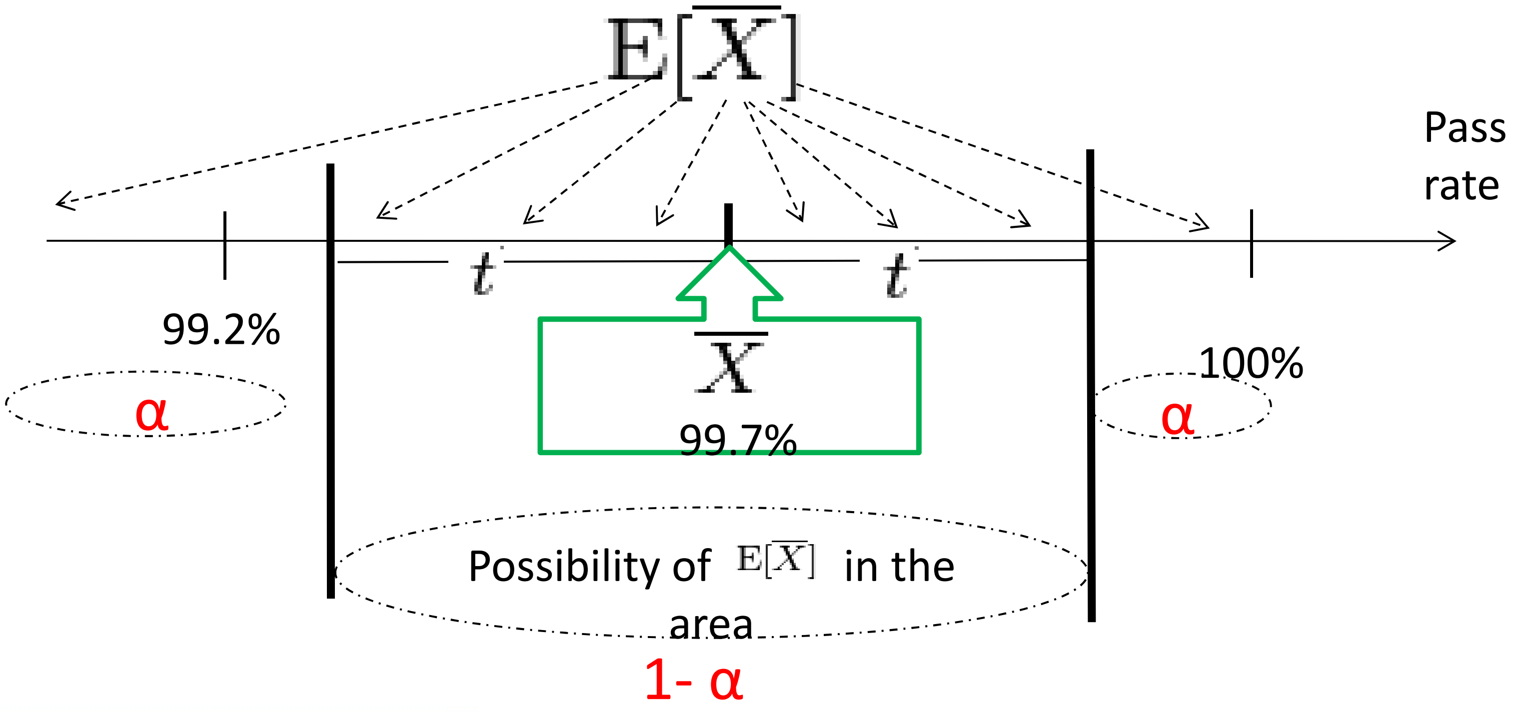
How **Possible** $E[\overline{X}]$ Close to $\overline{X}$?

Small α

Redefine-Coverage

Coverage = t AND α

High Coverage: $\mathrm{E}[\overline{X}]$ has high probability(1-a) close enough(t) to $\overline{X}$



- Way to Get Coverage

Pass rate in testing $\overline{X}$

Confidence Interval t

Confidence Level $1-\alpha$

?

$E[\overline{X}]$

ok in real use

Hoeffding Inequality

Probability Theory->Law of Large Numbers-> Hoeffding Inequality

A set of **Independent** Random Variable $X_1, X_2, X_3, \ldots,, X_N$

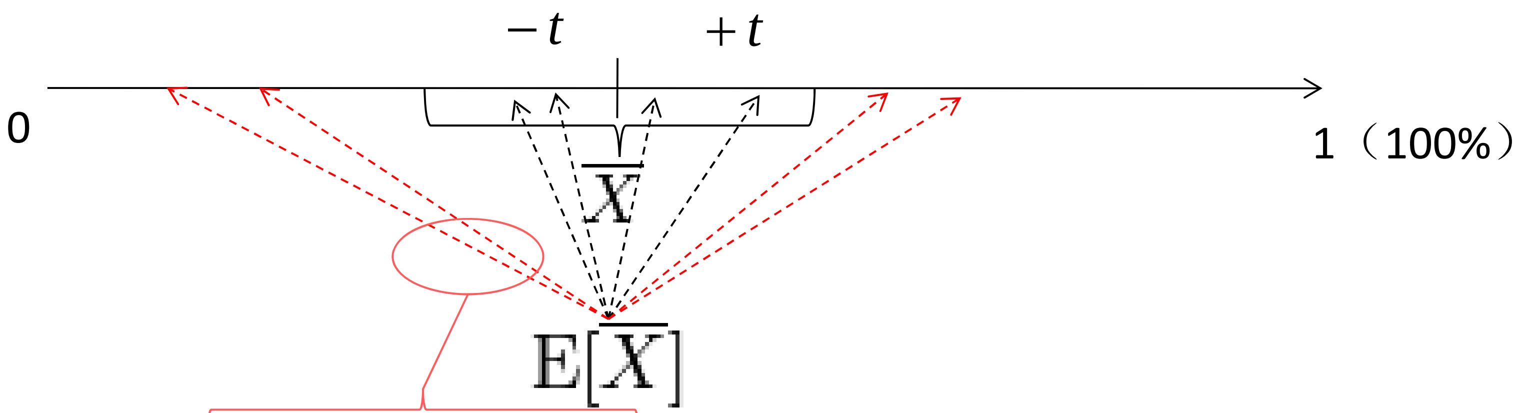
$$X_i \in [a, b] \qquad \overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$E(\overline{X})$ mathematical expectation

$$\boxed{\mathbb{P}(|\overline{X} - \mathrm{E}[\overline{X}]| \geq t) \leq 2e^{-2nt^2}}$$

Hoeffding Inequality

Probability Theory->Law of Large Numbers-> Hoeffding Inequality



$$\mathbb{P}(|\overline{X} - \mathrm{E}[\overline{X}]| \geq t) = \alpha \leq 2e^{-2nt^2}$$

Test result Real result interval Confidence Case Amount

Probability

Real result deviate test result larger than t

Probability of “Real result deviate test result larger than t”

- Way to Get Coverage

Pass rate in testing	$\overline{X}$
Confidence Interval	t
Confidence Level	1-α
Case Amount	n

$$\mathbb{P}(|\overline{X} - \mathrm{E}[\overline{X}]| \geq t) \leq 2e^{-2nt^2} \Rightarrow \mathrm{E}[\overline{X}]$$

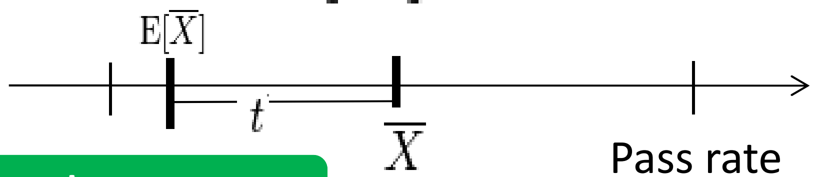
Hoeffding Inequality Utilization

Probability Theory->Law of Large Numbers-> Hoeffding Inequality

$$\mathbb{P}(|\overline{X} - \mathrm{E}[\overline{X}]| \geq t) = \alpha \leq 2e^{-2nt^2}$$

$$t \leq \sqrt{-\frac{\ln(a/2)}{2n}}$$

smallest $\mathrm{E}[\overline{X}] = \overline{X} - t$



Utilization 1

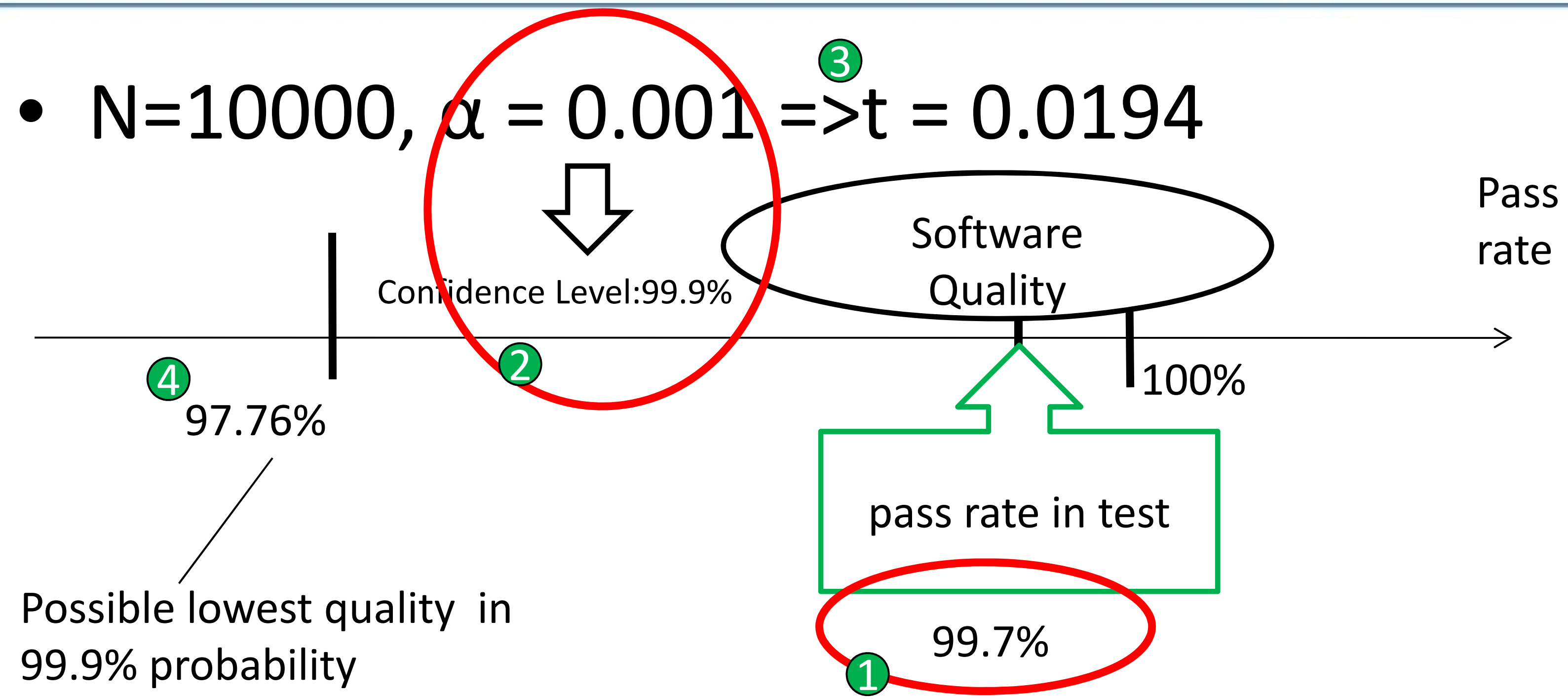
Estimate smallest $\mathrm{E}[\overline{X}]$

$$n \geq -\frac{\ln(a/2)}{2t^2}$$

Utilization 2

Estimate case amount

- N=10000, α = 0.001 =>t = 0.0194



① Have the test result in total

② Consider a reasonable Confidence Level

③ Calculate t by Hoeffding

$$t \le \sqrt{-\frac{\ln(a/2)}{2n}}$$

④ Calculate lowest ok in real use

Utilization 1 : Decision Before Release to Market

- A Web Service
  - No risk about personal injury
  - There are risk about lost biz and income
  - There is operation to recover
- Total Test Amount is 6000 and 40000
- Confidence Level : 1-a =99%

Remain risk = “known risk in test” + “estimated risk left”

- Known risk in test : $(1-\overline{X})$
- Estimated risk left: $[1-(\overline{X}-t)]*a$

Remain risk = $(1-\overline{X})+[1-(\overline{X}-t)]*a$

Utilization 1 : Decision Before Release to Market

Check Points (K)	Test Cost (10'000$)	Confidence Level	Confidence Interval	Pass Rate	Lowest Real Pass Rate	Remain Risk Possibility	Risk Lost (10'000 $)	Lost Discount to Present (10'000 $)
6	0.2*6 ≈1. 2	99%	0.02	75.9 %	73.9% (75.9% -0.02= 73.9%)	≤0. 241+0.261*0.01=0.2436	1000	243
6	0.2*6 ≈1. 2	99%	0.02	99.9 %	97.9%	≤0. 001+0.021*0.01=0.00121	1000	1.21
40	0.2* 40 ≈8	99%	0.008	99.9 %	99.1%	≤0. 001+0. 009*0. 01= 0. 00109	1000	1.09

2000 $ per K cases)

1-75.9%=0.241; 1-73.9%=0.261; 1-99%=0.01

Keep Testing?

Keep Testing?

Keep Testing?

In Test Planning, estimate how many test cases needed by set α and t

$$\mathbb{P}(|\overline{X} - \mathrm{E}[\overline{X}]| \geq t) = \alpha \leq 2e^{-2nt^2}$$

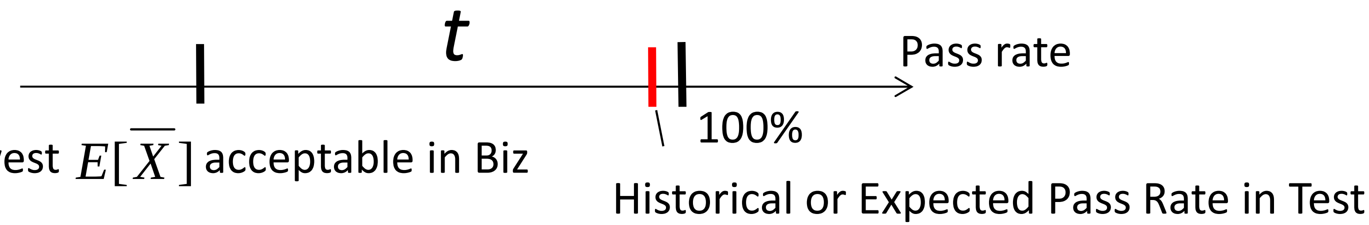
$$n \geq -\frac{\ln(a/2)}{2t^2}$$

$\mathrm{E}[\overline{X}]$ ok in real use

$\overline{X}$ Pass rate in testing

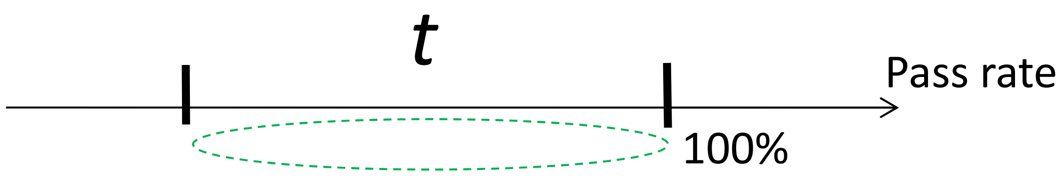
Utilization 2 : Test Estimation

1st set t by Business Risk:



Web Service as Utilization 1, set t = 0.005 (99.5% in real when pass rate =100%)

2nd set α by acceptable low possibility ($|E[\overline{X}]-\overline{X}|>t$ in low possibility)



Web Service as Utilization 1, set 1-α >99%

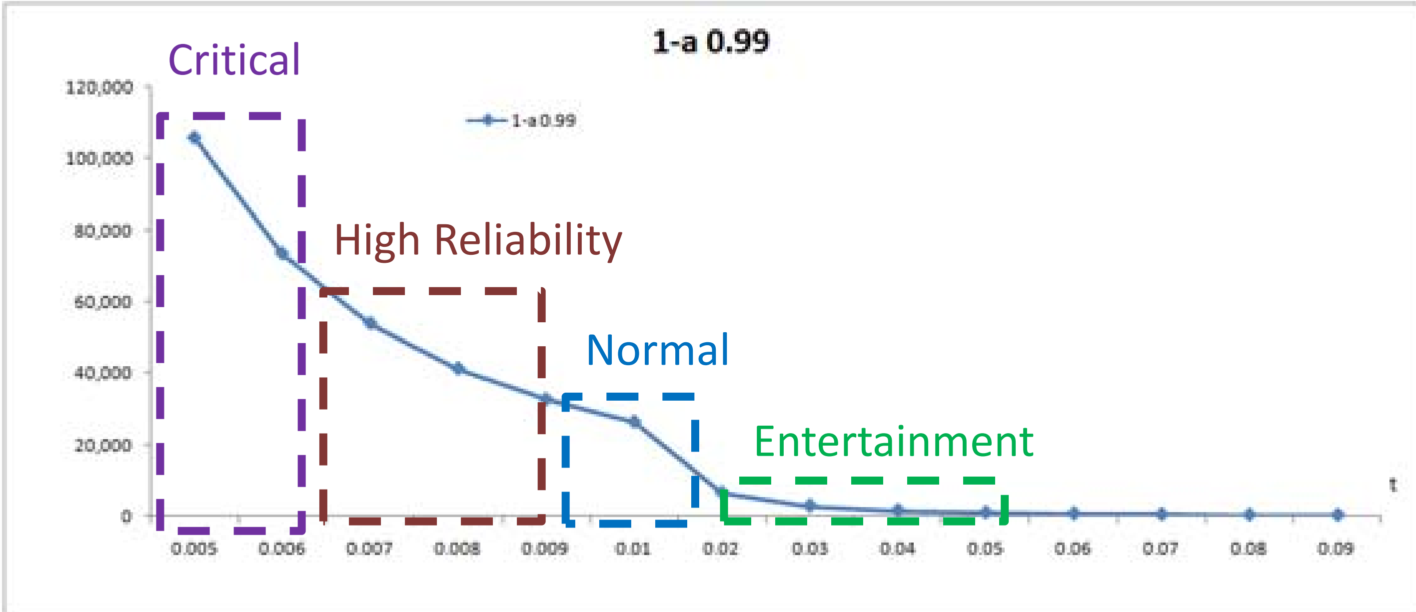
3rd Calculate n

Utilization 2 : Test Estimation

$$n \geq -\frac{\ln(a/2)}{2t^2}$$

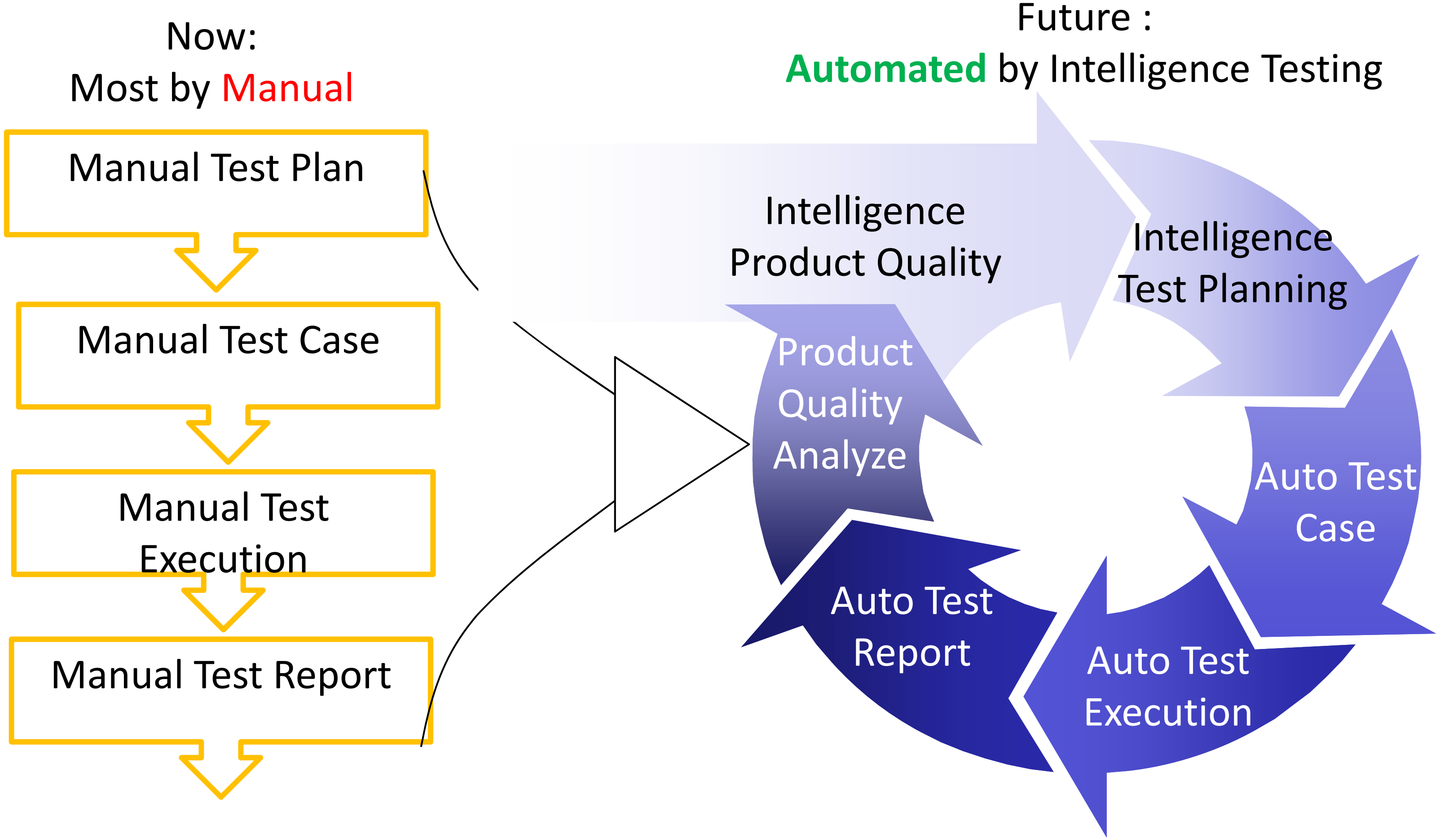
t = 0.005 1-α >99% → N=106000

- when α and t is small, n can be very big
- α = 0.001, t = 0.001 => n = 3’800’451

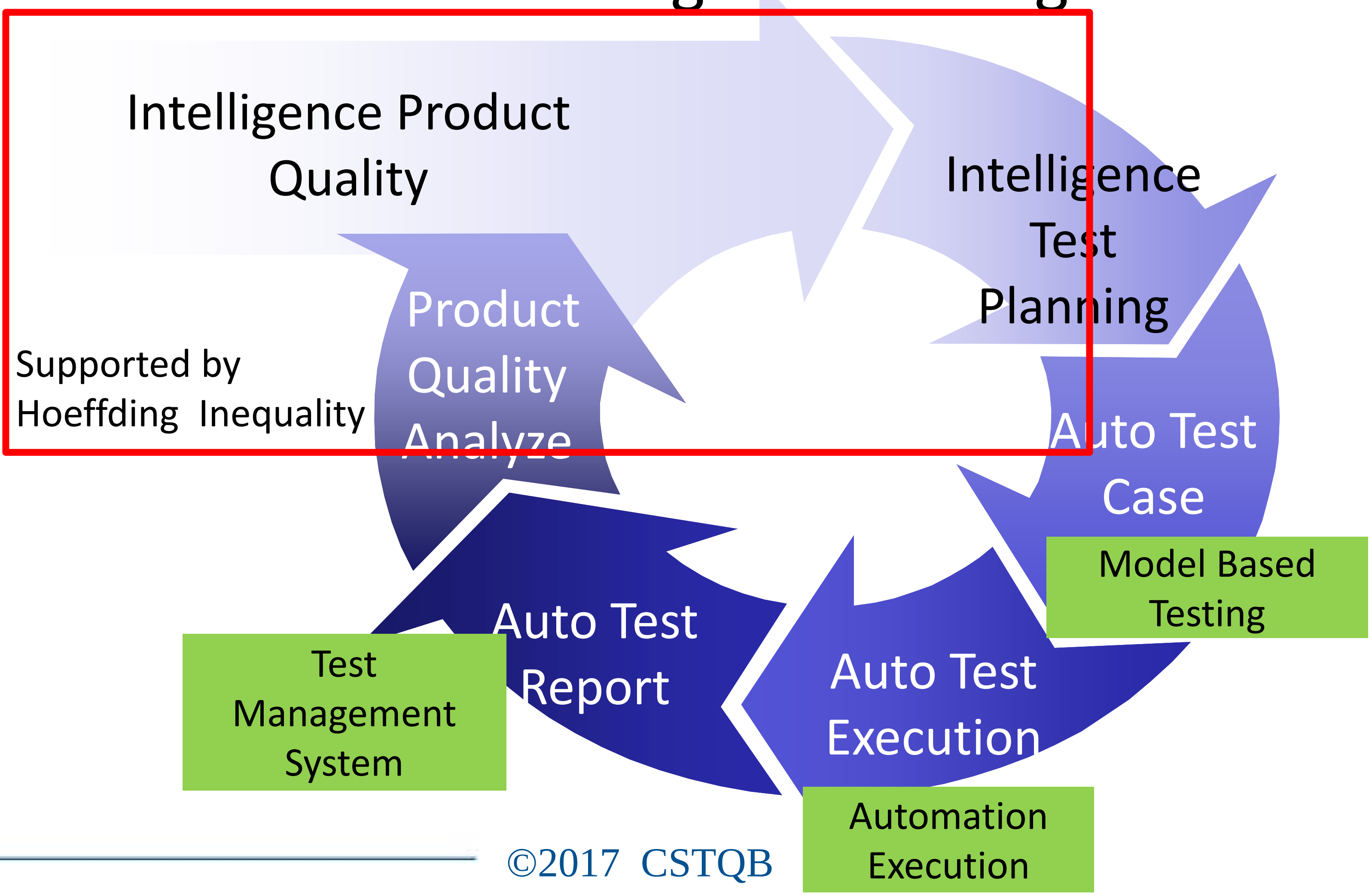


Future Application: Intelligence Testing

- Be the base of "intelligence testing"

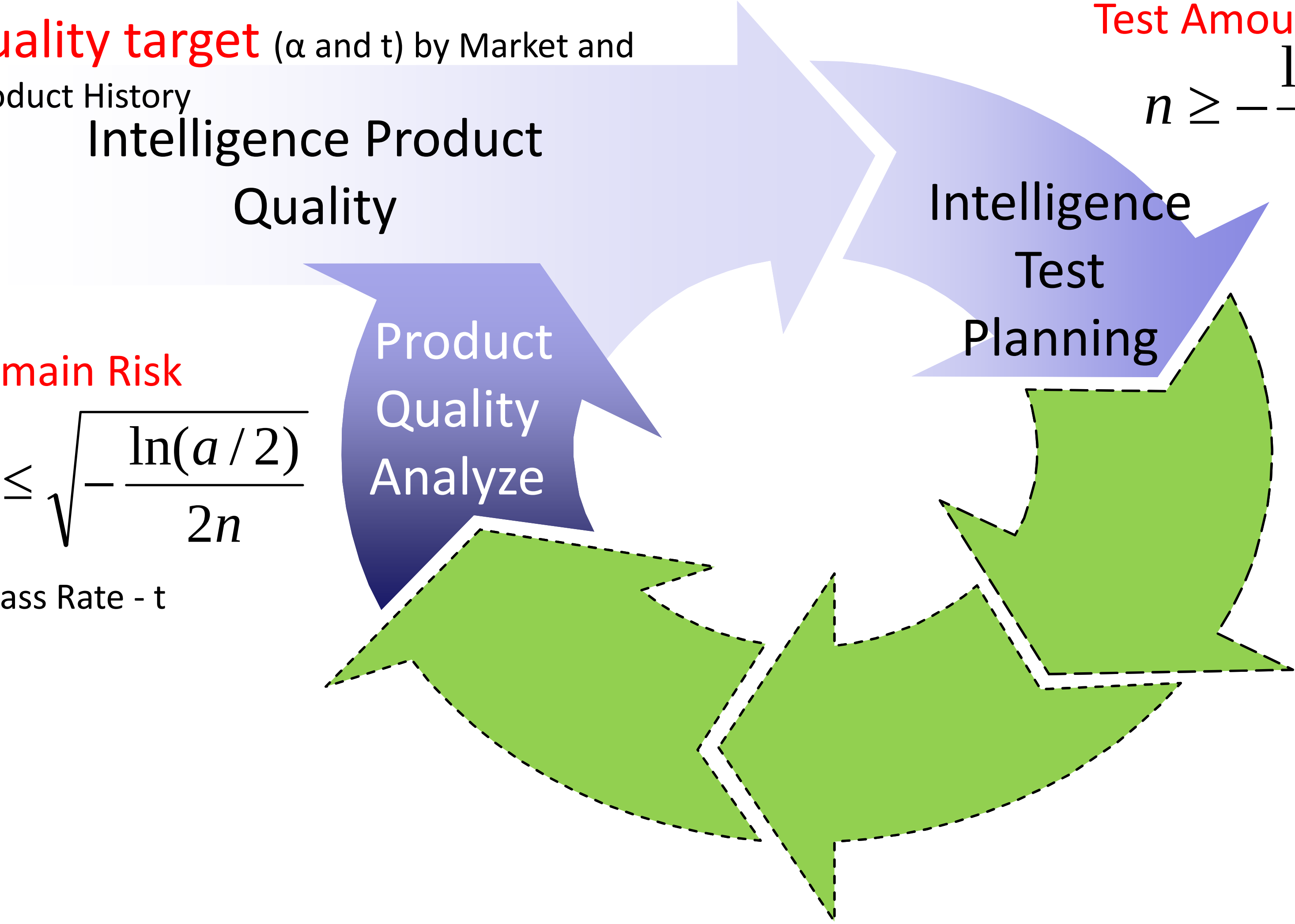


- Be the base of “intelligence testing”



Future Application: Intelligence Testing

quality target (α and t) by Market and Product History



Test Amount

$$n \geq -\frac{\ln(a/2)}{2t^2}$$

Remain Risk

$$t \leq \sqrt{-\frac{\ln(a/2)}{2n}}$$

=Pass Rate - t

Contact

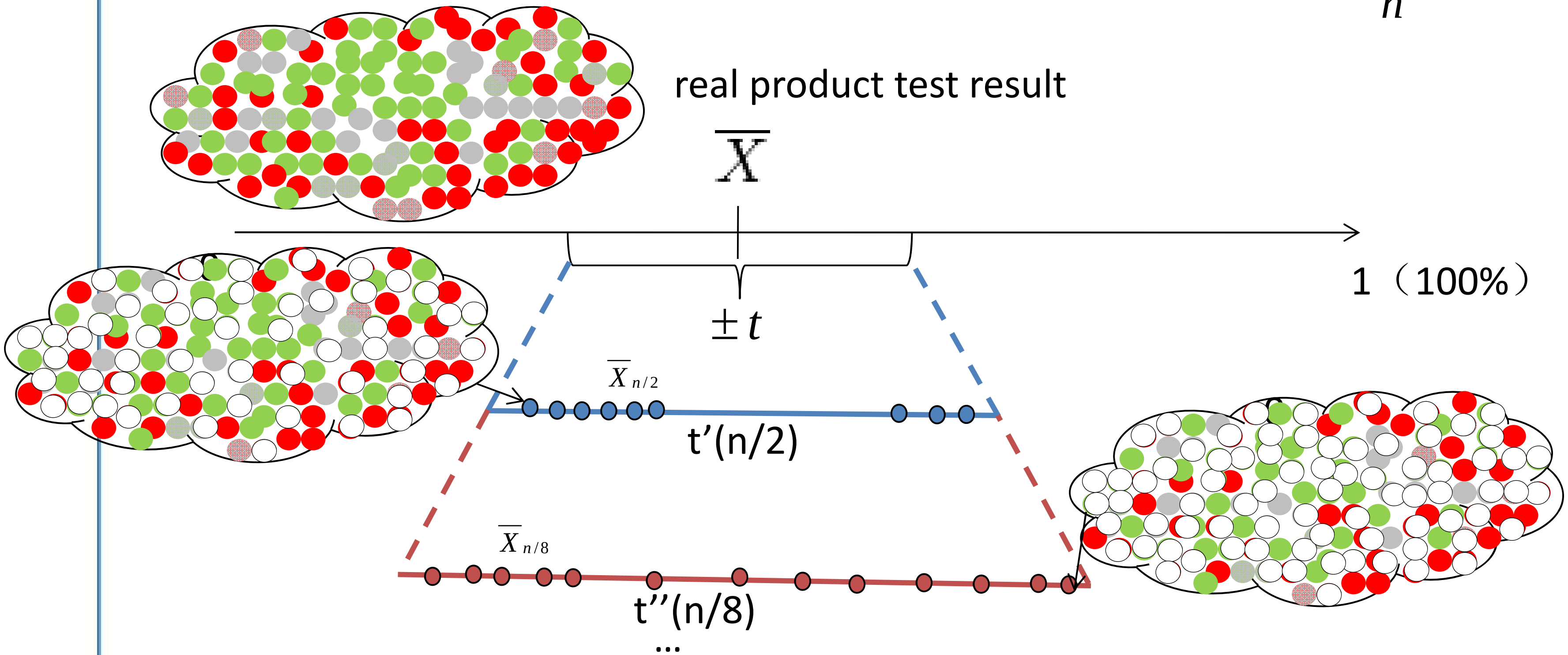
- More real project Data analyze
- Anyone who interested in further study, contact me by : geng_chen619@163.com

Thank you

ISTQB®让测试更专业

Appendix: Proven Real Project Data

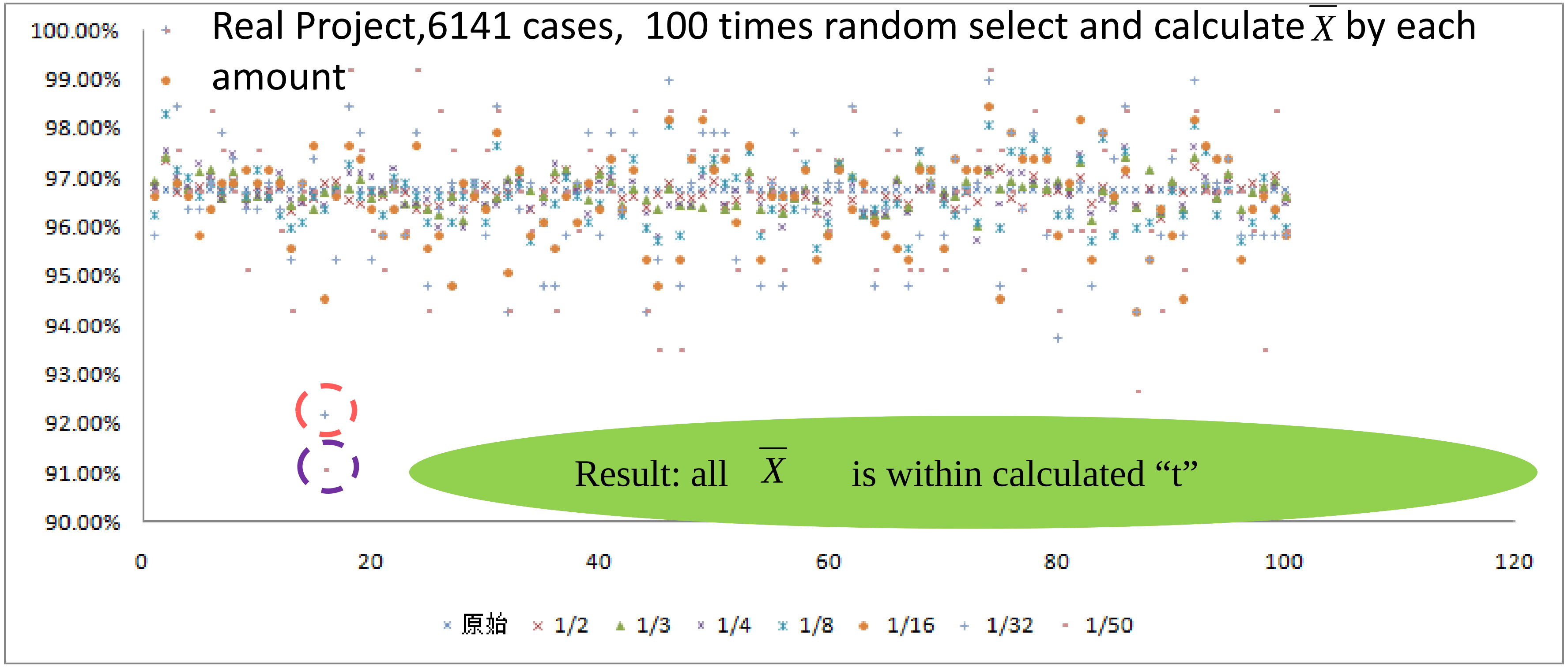
- With n, defined α , $\overline{X}$ = pass rate = $\frac{passedcases}{n}$



randomly select, calculate pass rate($\overline{X}$) , do 100 times, see the $\overline{X}$ is within boundary or not

$\overline{X}$ Pass rate in testing

Appendix: Proven Real Project Data



1/32, lower boundary 84.98% (t=0.117)

1/50, lower boundary 82.05% (t=0.146)

also verified in real project, 31000+ cases