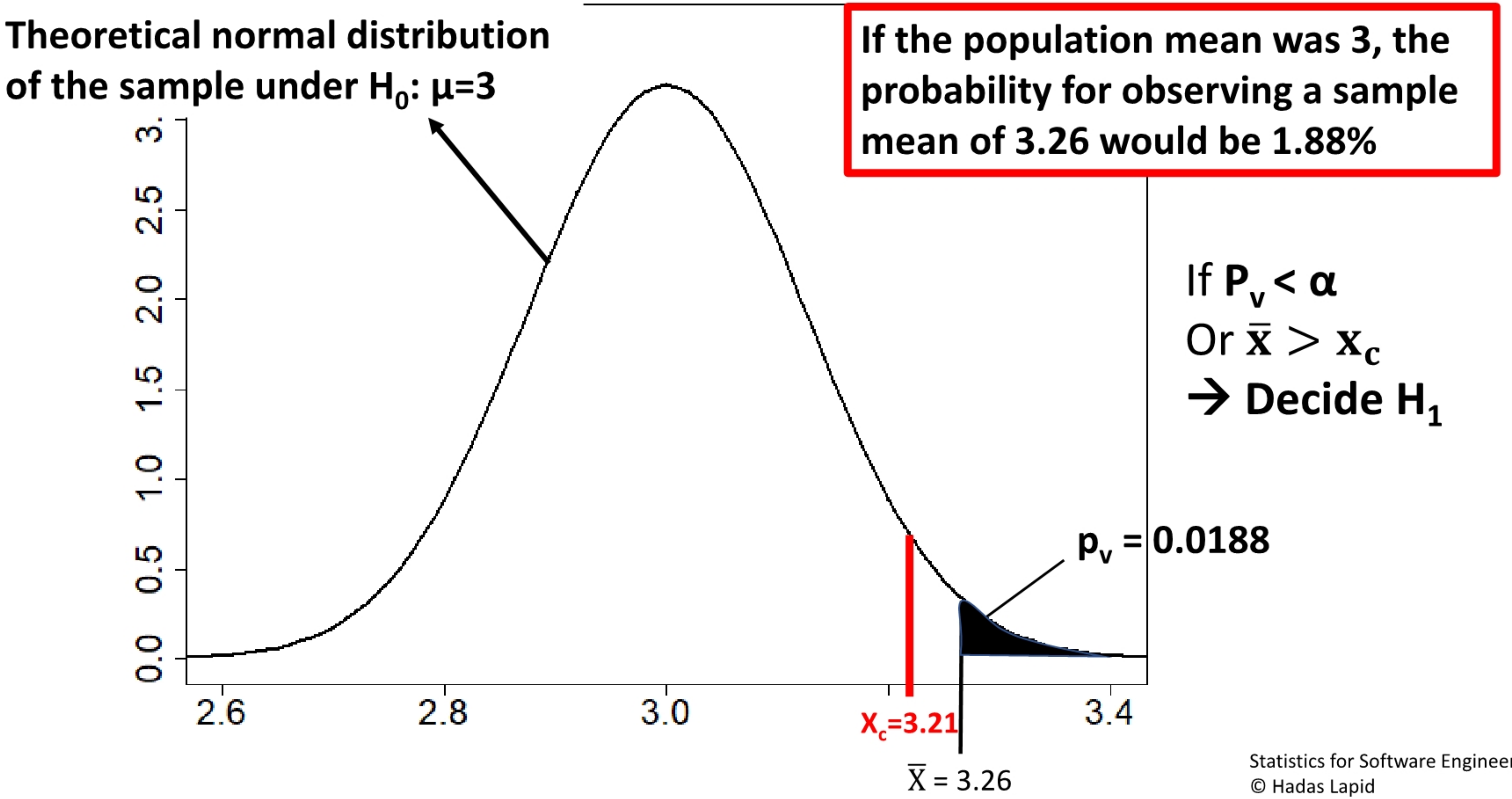
Statistical Methodology for Software Engineering

Hadas Lapid, PhD

Contents

- 1-sided test recap
- 2-sided hypothesis testing
- Goodness of fit - Chi-squared test
  - theory
  - example

1-sided Z test



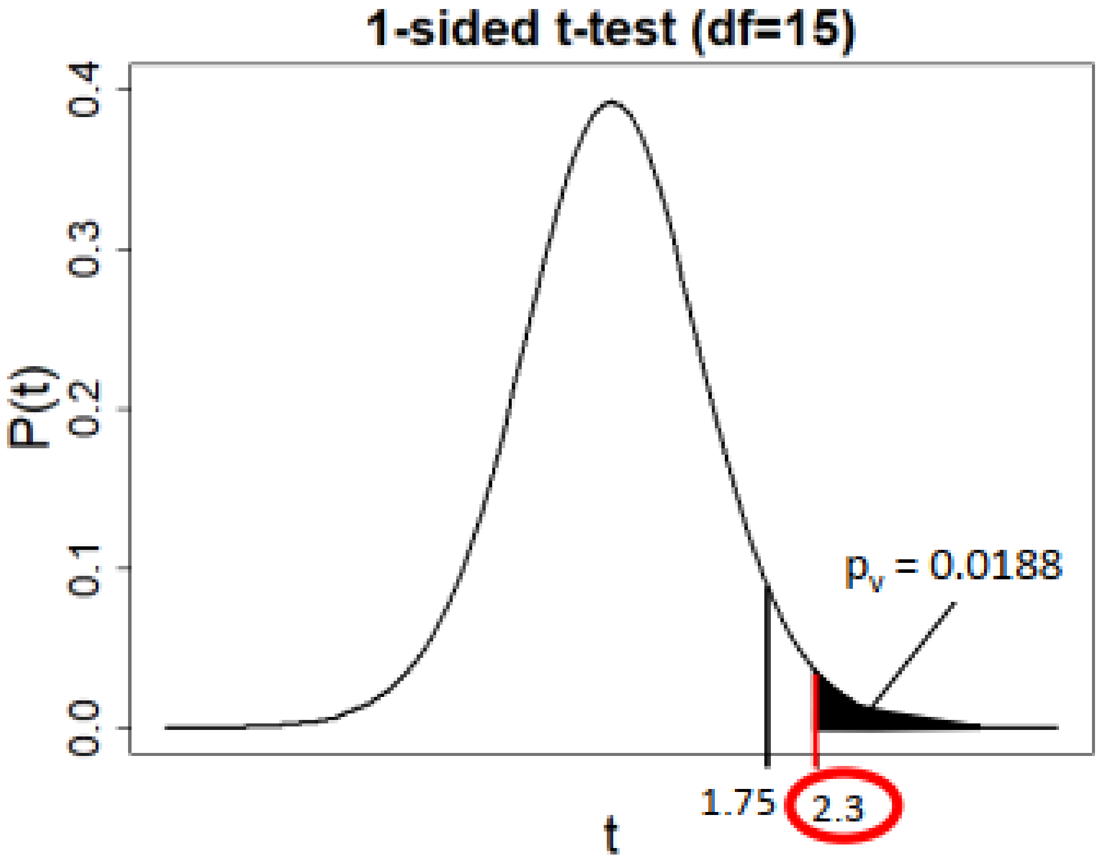
If $P_v < \alpha$

Or $\bar{x} > x_c$

→ Decide H_1

1-sided t-test

In case the population s.d., σ, is unknown – use t(df) distribution



t-statistic = 2.3

P_v = 0.0188

$\mathbf{t_{0.95}}$ = 1.75

2.3 > 1.75

→ Decide H_1

t-statistic is the t value associated with the P v area derived from $\bar{x}$

2-Sided hypothesis testing

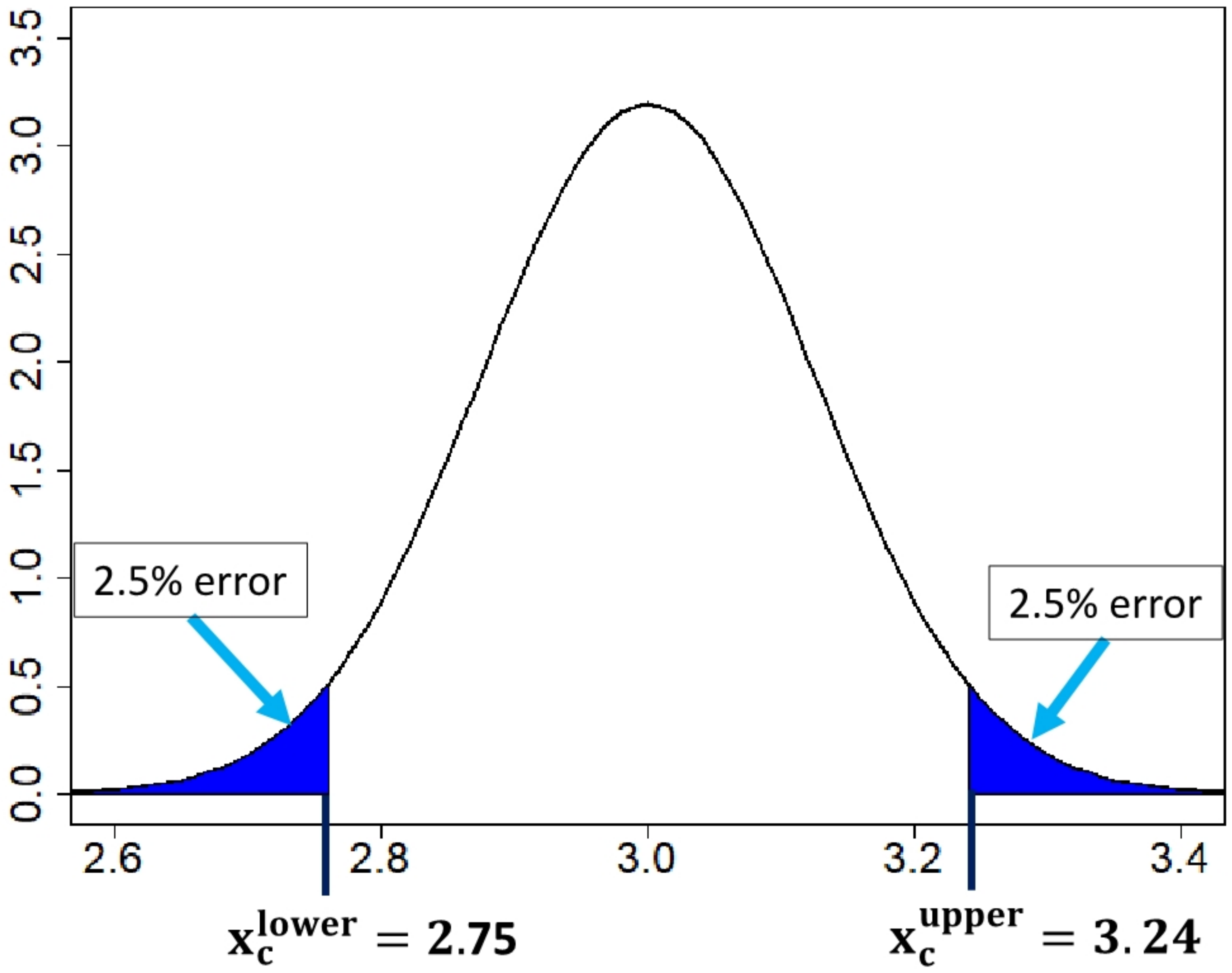
- Nonspecific directional hypothesis:

$$H_0:\ \mu = 3$$
$$H_1:\ \mu \neq 3$$

- Rejection region includes 2-sided extremes

$$|\bar{x} - \mu| > Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\mathrm{X_c} = \mu \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$



- For a given $\bar{x}$, p-value is the sum of the 2-tailed errors.

One-sample z/t-test summary

- When population sd (σ) is unknown:

 - use sample sd, S, and the **t distribution** with df=n-1.

 - such hypothesis test for μ is called **One-sided t-test**

- t($P = P_v$) is called the **test statistic**

- **2-sided test (z or t)** for μ at significance level α is equivalent to confidence interval at 1-α confidence level

Goodness of fit-motivation

Q: Is the sample distribution the same as the population one?

Canadian *population* blood types distribution

	Expected (%)
O	46%
A	42%
B	9%
AB	3%

Sample blood types distribution (n=465)

		Expected (%)
O	177	38%
A	187	40%
B	74	16%
AB	27	6%

Pearson's chi-squared test

Based on the central limit theorem,

Supposed n observations classified in *k* mutually exclusive groups: $x_i \in \{x_1, x_2 \dots x_k\}$

H_0: $m_i = n \cdot P_i \; \forall \; i$ (m_i observations for the *i*th class with probability P_i)
H_1: $m_i \neq n \cdot P_i \; \forall \; i$

$$\sum_{i=1}^{k} P_i = 1$$

$$n = \sum_{i=1}^{k} m_i = n \cdot \sum_{i=1}^{k} P_i = \sum_{i=1}^{k} x_i$$

Define X^2 (the sum of squared errors)

$$\boxed{X^2} = \sum_{i=1}^{k} \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} = \sum_{i=1}^{k} \frac{(x_i - m_i)^2}{m_i} = \boxed{\sum_{i=1}^{k} \frac{x_i^2}{m_i} - n}$$

Under H_0, as $n \to \infty$, x^2 distributes as χ^2 **(chi-squared)**

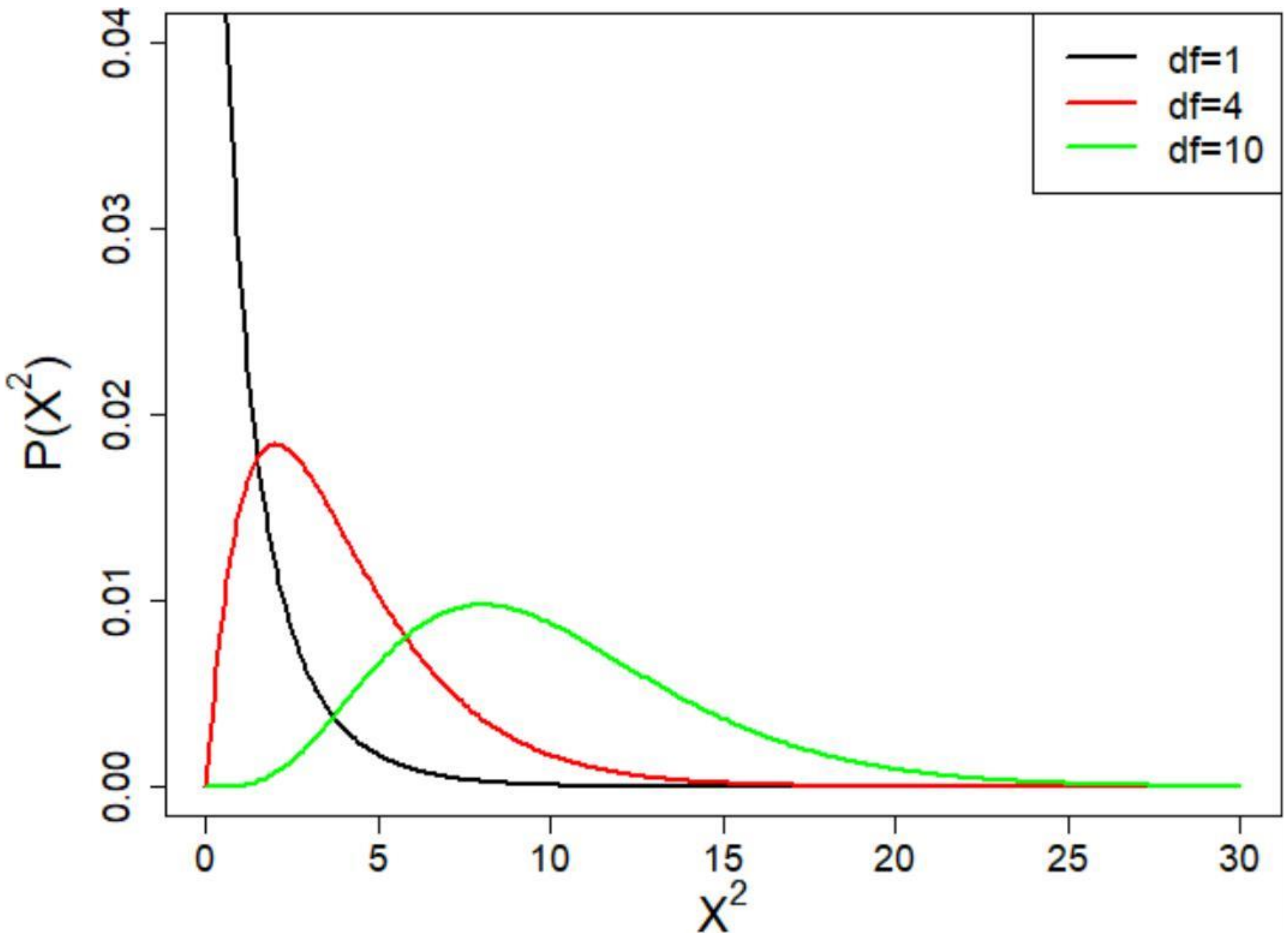
Calculating chi-square degrees of freedom

	Canadian distribution	***Sample distribution***
O	46%	38%
A	42%	40%
B	9%	16%
AB	3%	6%

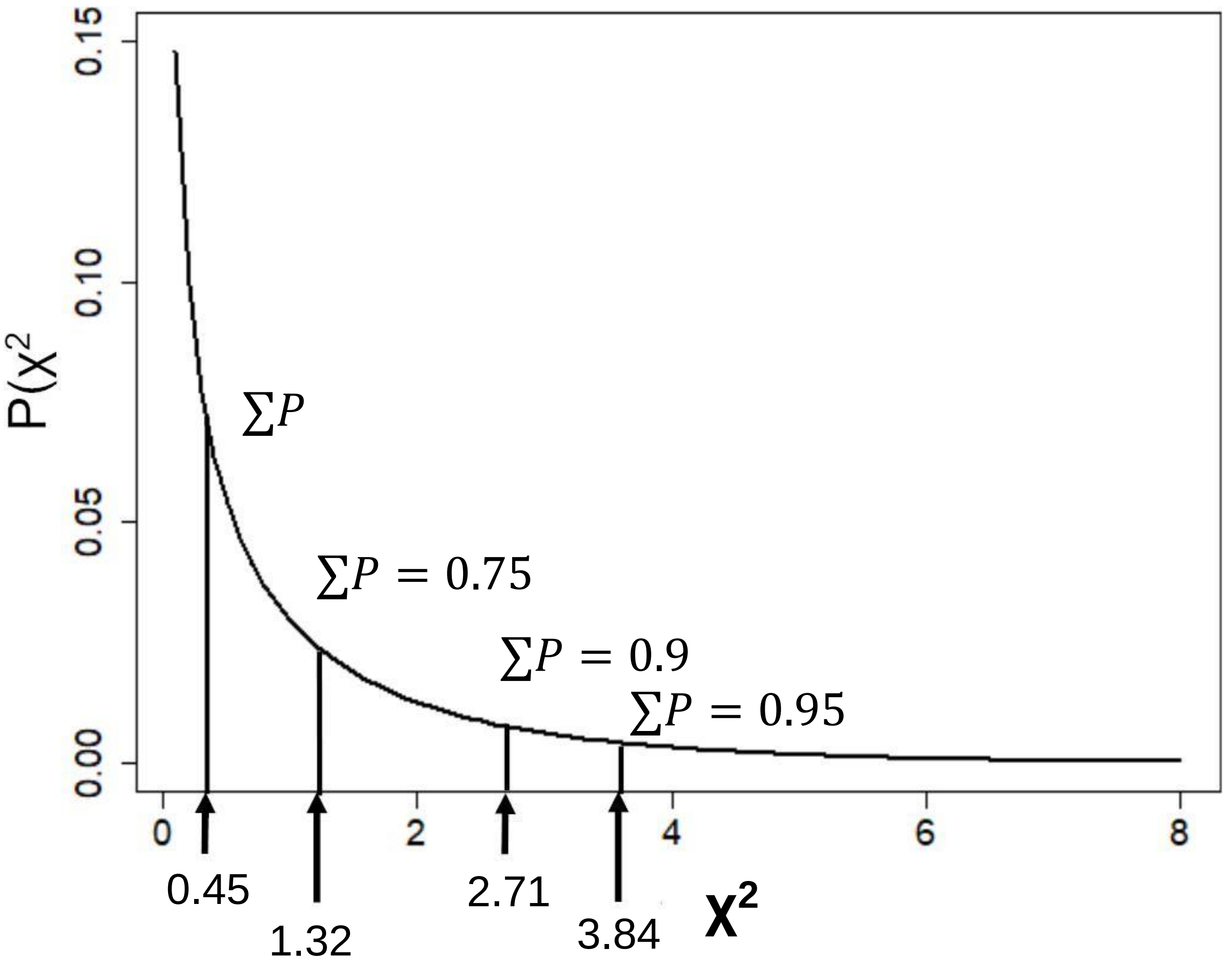
df = (number of groups-1) X (number of alternative distributions-1)

= (4-1) X (2-1) = 3

Chi-squared distributions



Chi-squared for df=1 critical values



Chi-squared test – numeric example

n=465	P Expected (%)	m_i (counts expected)	P Observed (%)	x_i (counts observed)	$\chi^2 = \frac{(\text{Observed-Expected})^2}{\text{Expected}}$
O	46%	0.46*465=213.9	38%	177	$(177 - 213.9)^2 / 213.9 = 6.366$
A	42%	0.42*465=195.3	40%	187	$(187 - 195.3)^2 / 195.3 = 0.353$
B	9%	0.09*465=41.85	16%	74	$(74 - 41.85)^2 / 41.85 = 24.698$
AB	3%	0.03*465=13.95	6%	27	(27 - 13.95)2 / 13.95 = 12.208
SUM	100%	465	100%	465	43.625

$\alpha=0.05$

$\chi_c^2 = X^2(0.95,df=3)$

= qchisq(0.95,df=3) = 7.815

Alternatively, chisq.test(x_i,p=p_exp)

X^2(observed) = 43.625 > χ_c^2

P_V = pchisq(43.625,df=3,lower.tail=FALSE)

= 1.8e-9 << 0.05

→ Reject H_0