

# **M00C Statistique pour ingénieur**

## **Thème 3 : tests d'hypothèses, analyse de la variance**

### **Vidéo 4 : Test d'ajustement et test d'indépendance**

Martial Sauceau

Mines Albi

# Les tests d'hypothèses

**Test d'ajustement à une loi**

**Test d'indépendance de deux variables qualitatives**

# Test d'ajustement à une loi

Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi
Effectifs observés	30	20	41	20	39

Répartition uniforme des pannes selon les 5 jours de la semaine ?

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$\mathcal{H}_1$  : la répartition n'est pas uniforme

# Test d'ajustement à une loi

$N_i$	Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi	n
	Effectifs observés	30	20	41	20	39	150

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$p_i = 1/5 \qquad n_i = n \times p_i = 30$$

# Test d'ajustement à une loi

	Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi	n
$N_i$	Effectifs observés	30	20	41	20	39	150
$n \times p_i$	Effectifs théoriques	30	30	30	30	30	150

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$d^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} \sim \chi^2(k - 1 - l) = \chi^2(4)$$

nombre de classes

nombre de paramètres estimés

# Test d'ajustement à une loi

	Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi	n
$N_i$	Effectifs observés	30	20	41	20	39	150
$n \times p_i$	Effectifs théoriques	30	30	30	30	30	150

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$d^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} \sim \chi^2(k - 1 - l) = \chi^2(4)$$

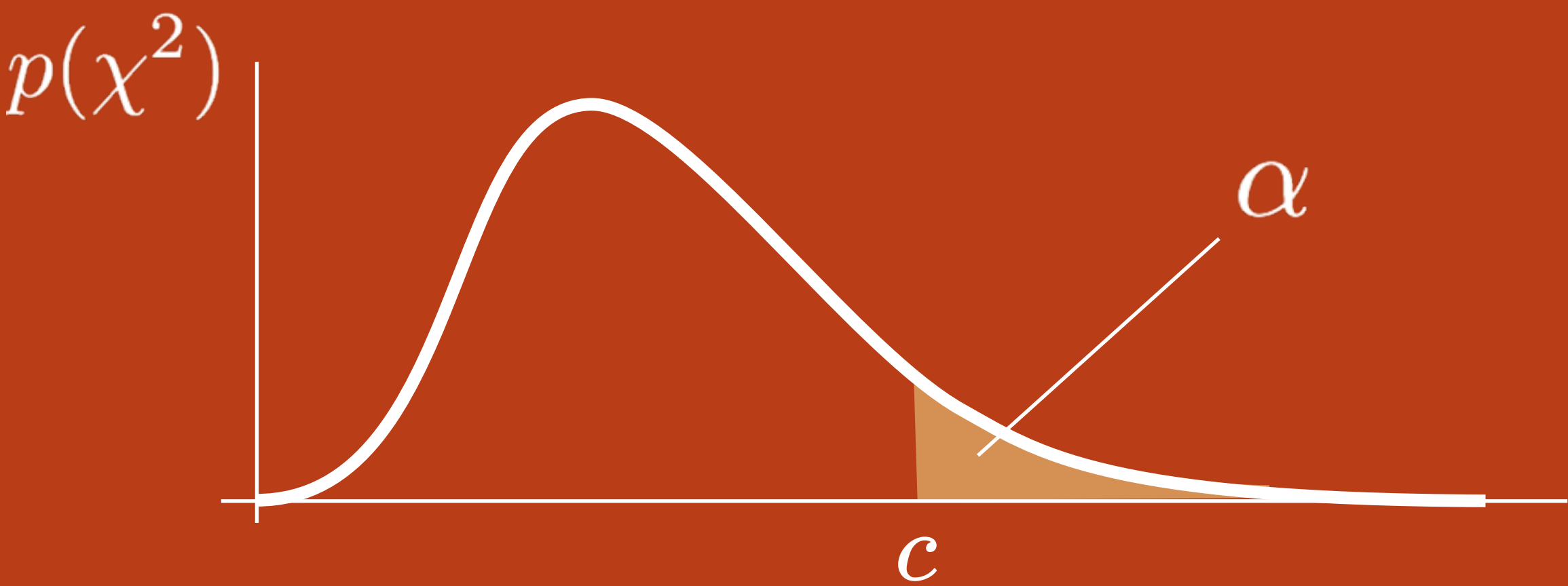
$$n \times p_i > 5$$

# Test d'ajustement à une loi

	Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi	n
$N_i$	Effectifs observés	30	20	41	20	39	150
$n \times p_i$	Effectifs théoriques	30	30	30	30	30	150

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$d^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i}$$

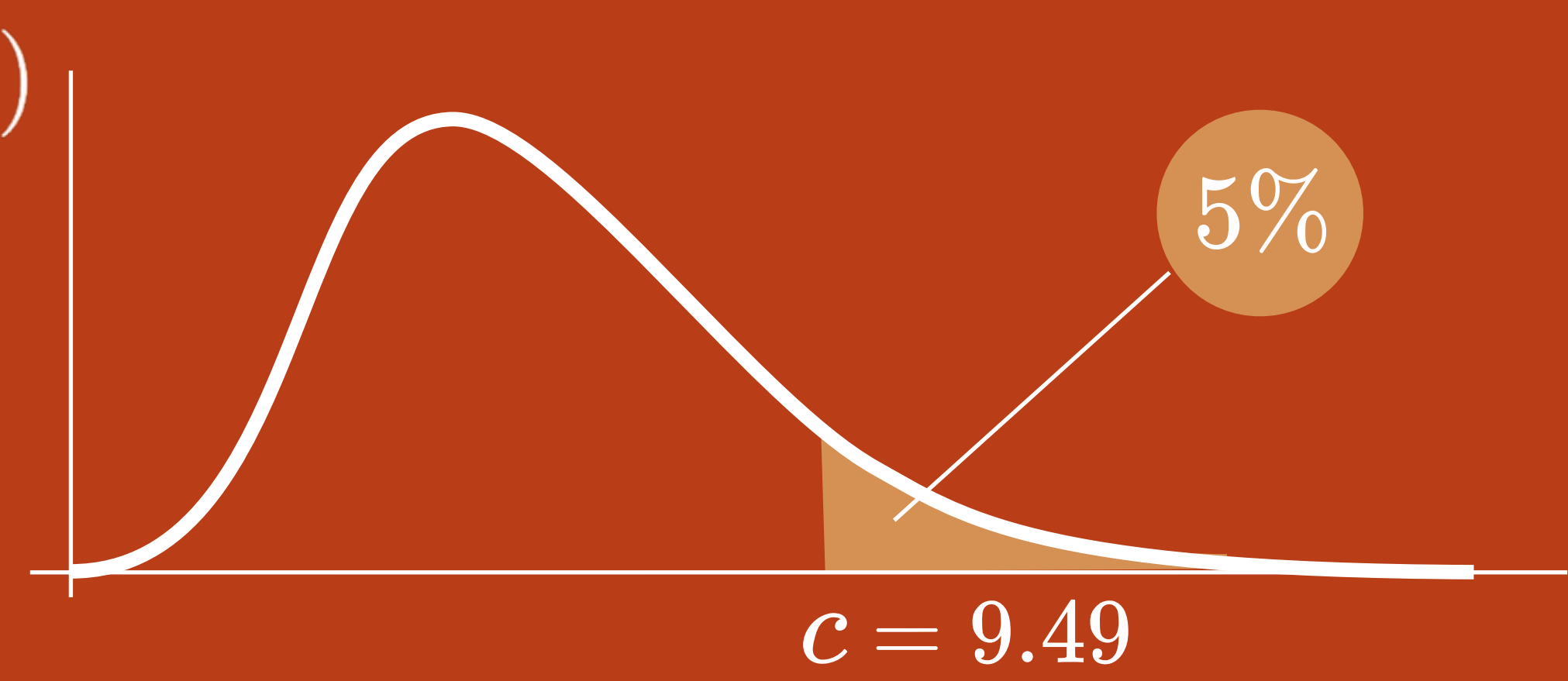


# Test d'ajustement à une loi

<b>N<sub>i</sub></b>  <b>n x p<sub>i</sub></b>	Jour X <sub>i</sub>	lundi	mardi	mercredi	jeudi	vendredi	n
	Effectifs observés	30	20	41	20	39	150
	Effectifs théoriques	30	30	30	30	30	150
	Distance	0	3.33	4.03	3.33	2.70	13.4

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$d^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} = 13.4 > c \Rightarrow \cancel{\mathcal{H}_0}$$





# Test d'ajustement à une loi

<div> <div><math>N_i</math></div> <div><math>n \times p_i</math></div> </div>	Jour $X_i$	lundi	mardi	mercredi	jeudi	vendredi	n
	Effectifs observés	30	20	41	20	39	150
	Effectifs théoriques	30	30	30	30	30	150
	Distance	0	3.33	4.03	3.33	2.70	13.4

$\mathcal{H}_0$  : la répartition des pannes est uniforme

$$d^2 = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} = 13.4 > c$$

$\Rightarrow \mathcal{H}_0$

la répartition des pannes pas uniforme

p-valeur = 0.009

# Les tests d'hypothèses

Test d'ajustement à une loi

Test d'indépendance de deux variables qualitatives

# Test d'indépendance de deux variables qualitatives

A \ B	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	52	41	93
Yeux bruns	98	190	288
Yeux verts	26	62	88
$\Sigma$	176	293	469

A \ B	B <sub>1</sub>	B <sub>2</sub>	$\Sigma$
A <sub>1</sub>	n <sub>11</sub>	n <sub>12</sub>	n <sub>1.</sub>
A <sub>2</sub>	n <sub>21</sub>	n <sub>22</sub>	n <sub>2.</sub>
A <sub>3</sub>	n <sub>31</sub>	n <sub>32</sub>	n <sub>3.</sub>
$\Sigma$	n <sub>.1</sub>	n <sub>.2</sub>	n

La couleur des yeux est-elle liée à la couleur des cheveux ?

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$\mathcal{H}_1$  : la couleur des yeux et celle des cheveux ne sont pas indépendantes

# Test d'indépendance de deux variables qualitatives

A \ B	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	52	41	93
Yeux bruns	98	190	288
Yeux verts	26	62	88
$\Sigma$	176	293	469

A \ B	B <sub>1</sub>	B <sub>2</sub>	$\Sigma$
A <sub>1</sub>	n <sub>11</sub>	n <sub>12</sub>	n <sub>1.</sub>
A <sub>2</sub>	n <sub>21</sub>	n <sub>22</sub>	n <sub>2.</sub>
A <sub>3</sub>	n <sub>31</sub>	n <sub>32</sub>	n <sub>3.</sub>
$\Sigma$	n <sub>.1</sub>	n <sub>.2</sub>	n

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$$p_{ij} = p_{i.} \times p_{.j} \quad \hat{p}_{i.} = \frac{n_{i.}}{n} \quad \hat{p}_{.j} = \frac{n_{.j}}{n} \quad \hat{p}_{ij} = \frac{n_{ij}}{n}$$

# Test d'indépendance de deux variables qualitatives

A \ B	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	52	41	93
Yeux bruns	98	190	288
Yeux verts	26	62	88
$\Sigma$	176	293	469

A \ B	B <sub>1</sub>	B <sub>2</sub>	$\Sigma$
A <sub>1</sub>	n <sub>11</sub>	n <sub>12</sub>	n <sub>1.</sub>
A <sub>2</sub>	n <sub>21</sub>	n <sub>22</sub>	n <sub>2.</sub>
A <sub>3</sub>	n <sub>31</sub>	n <sub>32</sub>	n <sub>3.</sub>
$\Sigma$	n. <sub>1</sub>	n. <sub>2</sub>	n

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$$p_{ij} = p_{i.} \times p_{.j} \qquad \hat{p}_{ij} = \frac{n_{i.} \times n_{.j}}{n \times n} \qquad t_{ij} = \frac{n_{i.} \times n_{.j}}{n}$$

# Test d'indépendance de deux variables qualitatives

$n_{ij}$	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	52	41	93
Yeux bruns	98	190	288
Yeux verts	26	62	88
$\Sigma$	176	293	469

$t_{ij}$	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	34,9	58,1	93
Yeux bruns	108,1	179,9	288
Yeux verts	33,0	55,0	88
$\Sigma$	176	293	469

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$$p_{ij} = p_{i.} \times p_{.j} \quad \hat{p}_{ij} = \frac{n_{i.} \times n_{.j}}{n \times n}$$

$$t_{ij} = \frac{n_{i.} \times n_{.j}}{n}$$

# Test d'indépendance de deux variables qualitatives

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

r modalités	s modalités		
	Cheveux clairs	Cheveux sombres	$\Sigma$
	Yeux bleus		
	Yeux bruns		
	Yeux verts		
	$\Sigma$		

$$d^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - t_{ij})^2}{t_{ij}} \sim \chi^2((r-1)(s-1)) \sim \chi^2(2)$$

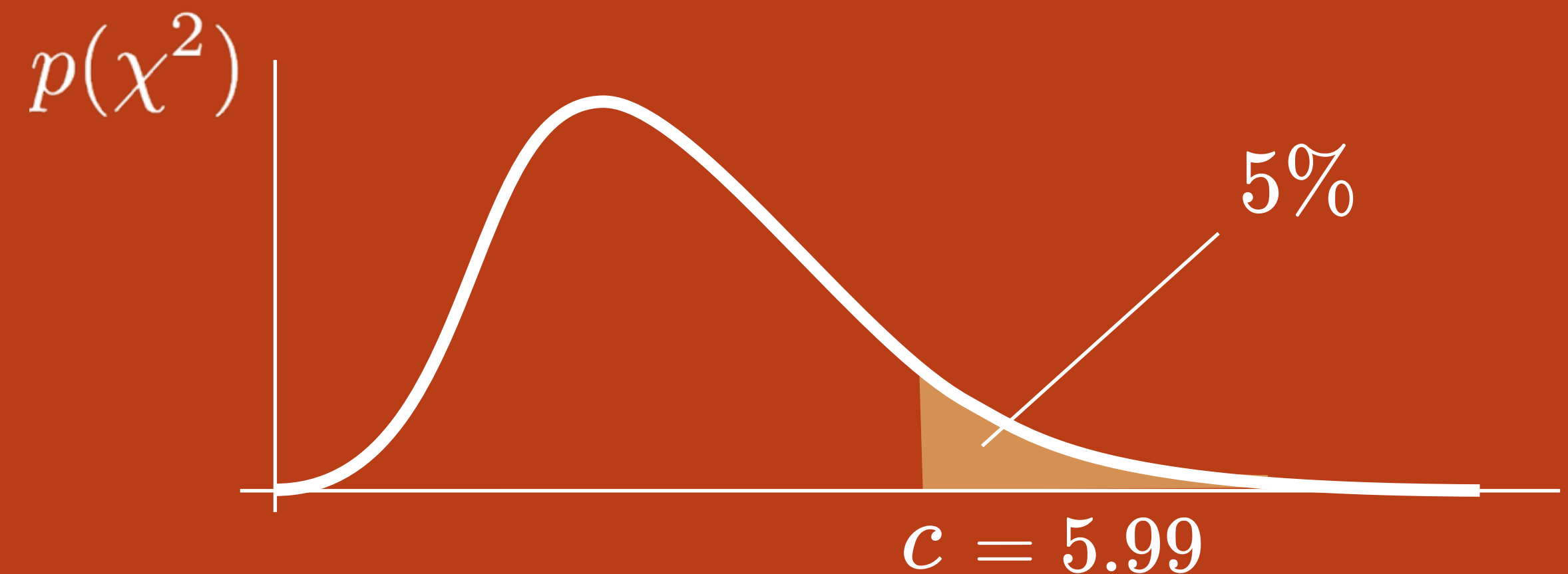
# Test d'indépendance de deux variables qualitatives

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$$d^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - t_{ij})^2}{t_{ij}} \sim \chi^2(2)$$

$$d^2 > c \Rightarrow \cancel{\mathcal{H}_0}$$

r modalités	s modalités		
	$\frac{(n_{ij} - t_{ij})^2}{t_{ij}}$	Cheveux clairs	Cheveux sombres
	Yeux bleus	8,38	5,03
	Yeux bruns	0,94	0,57
	Yeux verts	1,48	0,89
	$\Sigma$		17,3





# Test d'indépendance de deux variables qualitatives

$\mathcal{H}_0$  : la couleur des yeux et celle des cheveux sont indépendantes

$$d^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - t_{ij})^2}{t_{ij}} \sim \chi^2(2)$$

$$d^2 > c \Rightarrow \cancel{\mathcal{H}_0}$$

r modalités

s modalités

$\frac{(n_{ij} - t_{ij})^2}{t_{ij}}$	Cheveux clairs	Cheveux sombres	$\Sigma$
Yeux bleus	8,38	5,03	
Yeux bruns	0,94	0,57	
Yeux verts	1,48	0,89	
$\Sigma$			17,3

la couleur des cheveux et la couleur des yeux ne sont pas indépendantes

p-valeur = 0.00017



I N S T I T U T  
Mines-Télécom

---