EdX and its Members use cookies and other tracking technologies for performance, analytics, and marketing purposes. By using this website, you accept this use. Learn more about these technologies in the Privacy Policy. ×

Course › Unit 2 Foundation of Inference › Lecture 6: Introduction to Hypothesis Testing, and Type 1 and Type 2 Errors › 3. Statistical Model of a Two Sample Experiment

Currently enrolled in **Audit Track** (expires December 25, 2019) Upgrade ($300)

3. Statistical Model of a Two Sample Experiment

Preparation: Statistical Model of a Two Sample Experiment

2/2 points (graded)

The observed outcome of a statistical experiment consists of two samples:

$$X_1, X_2, \ldots X_n \overset{\text{i.i.d.}}{\sim} X \sim \text{Ber}\,(p_1)$$

$$Y_1, Y_2, \ldots Y_m \overset{\text{i.i.d.}}{\sim} Y \sim \text{Ber}\,(p_2)\,.$$

where in addition, X and Y are independent.

An associated statistical model is $\left(E, \{P_\theta\}_{\theta\in\Theta}\right)$ where E is the (smallest) sample space of the pair (X, Y), and P_θ, is the joint distribution of (X, Y) with parameter θ. Because X and Y are independent, their joint distribution is the product of their respective distributions.

Identify the sample space E and the parameter space Θ:

(Choose one per column.)

Sample space E: Parameter space Θ

○ $\{0,1\}$

● $\{0,1\} \times \{0,1\} = \{(0,0),(0,1),(1,0),(1,1)\}$

○ $(0,1)$

○ $(0,1) \times (0,1) \in \mathbb{R}^2$

✓

○ $\{0,1\}$

○ $\{0,1\} \times \{0,1\} = \{(0,0),(0,1),(1,0),(1,1)\}$

○ $(0,1)$

● $(0,1) \times (0,1) \in \mathbb{R}^2$

✓

Solution:

Since $X \sim$ Ber (p_1) and $Y \sim$ Ber (p_2), the pair (X,Y) takes value in the sample space $E = \{0,1\} \times \{0,1\} = \{(0,0),(0,1),(1,0),(1,1)\}$.

Since X, Y are independent, the joint distribution of (X,Y) is the product Ber (p_1) $\times$Ber (p_2). Hence, the family $\{P_\theta\}_{\theta\in\Theta}$ of joint distributions is parametrized by $\theta = (p_1, p_2)$ and the parameter space is

$$\Theta = \{(p_1,p_2) : p_1 \in (0,1), p_2 \in (0,1)\} = (0,1) \times (0,1) \in \mathbb{R}^2.$$

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem

Preparation: Statistical Model of a Two Sample Experiment II

2/2 points (graded)

Recall the statistical experiment from the lecture: to test whether boarding times by the Window-Middle-Aisle boarding method is shorter than boarding times by the rear-to-front method, we collect a sample of boarding times of each method. We model these boarding times as the following two sets of normal variables:

$$\begin{aligned} X_1, X_2, \ldots X_n \text{ are } i.i.d. \text{ copies of } X \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right) \qquad &\text{boarding times of rear-to-front} \\ Y_1, Y_2, \ldots Y_m \text{ are } i.i.d. \text{ copies of } Y \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right) \qquad &\text{boarding times of window-middle-aisle} \end{aligned}$$

where X and Y are also independent.

Let $\left(E, \{P_\theta\}_{\theta\in\Theta}\right)$ be the statistical model associated with this experiment where

- E is the sample space of the pair of random variables (X, Y);
- $\{P_\theta\}_{\theta\in\Theta}$ is the family of joint distributions of (X, Y).

For simplicity, **assume the two variances σ_1 and σ_2 are some known, fixed quantities σ_1^* and σ_2^***.

Choose a valid candidate for the parametrization θ, which describes the family of joint probability distributions of (X, Y).

○ $\mu_1 - \mu_2$

○ $\left(\mu_1, (\sigma_1)^2, \mu_2, (\sigma_2)^2\right)$ where $(\sigma_1)^2$ and $(\sigma_2)^2$ can each take on more than a single value

● (μ_1, μ_2)

○ (μ_2, μ_1)

✓

Which of the following are legitimate choice(s) of the parameter space Θ?
(Choose all that apply)

☐ $\Theta = \mathbb{R}$

☐ $\Theta = [0, \infty)$

☑ $\Theta = \mathbb{R}^2$

☑ $\Theta = [0, \infty) \times [0, \infty)$

✓

Solution:

Since X, Y are independent, the joint distribution of (X, Y) is the product $\mathcal{N}\left(\mu_1, (\sigma_1)^2\right) \times \mathcal{N}\left(\mu_2, (\sigma_2)^2\right)$

Since the variances σ_1 and σ_2 are fixed and known, the only parameters determining the joint distribution is μ_1 and μ_2. Hence, a choice of the parameter θ is the 2-dimensional vector $\begin{pmatrix} \mu_1 & \mu_2 \end{pmatrix}$. (We could also have chosen to construct the statistical model using the pair (Y, X) instead. The family of joint distributions in that case would be parametrized by $\begin{pmatrix} \mu_2 & \mu_1 \end{pmatrix}$).

This gives the parameter space

$$\Theta = \{(\mu_1, \mu_2) : \mu_1 \in \mathbb{R}, \mu_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

Because μ_1 and μ_2 model average boarding times, we can further restrict to

$$\Theta = \{(\mu_1, \mu_2) : \mu_1 \in [0, \infty), \mu_2 \in [0, \infty)\} = [0, \infty) \times [0, \infty)\}.$$

Submit You have used 1 of 2 attempts

ℹ Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Foundation of Inference:Lecture 6: Introduction to Hypothesis Testing, and Type 1 and Type 2 Errors / 3. Statistical Model of a Two Sample Experiment

Add a Post

Show all posts ▼ by recent activity ▼

There are no posts in this topic yet.

✖

Audit Access Expires Dec 24, 2019

You lose all access to this course, including your progress, on Dec 24, 2019.

Upgrade by Nov 4, 2019 to get unlimited access to the course as long as it exists on the site. **Upgrade now**

Learn About Verified Certificates

© All Rights Reserved