6.

Setup: All problems on this page use the following setup.

Let $\mathcal{X} = \{1, \ldots, k\}$ be a sample space of k possible outcomes of an experiment, and $(\mathcal{X}, \{P_\theta\}_{\theta \in \Theta})$ be a statistical model for this experiment.

Suppose that this model has d parameters, so that $\Theta \subset \mathbb{R}^d$.

Suppose that the Fisher information matrix $I(\theta)$ of this statistical model is nonsingular (i.e. invertible) and that the MLE $\hat{\theta}_n$ has the Gaussian asymptotic distribution

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{\text{in } (d) \text{ under } P_\theta} \mathcal{N}\left(0, I^{-1}(\theta)\right)$$

for every $\theta \in \Theta$.

Let $X_1, \ldots, X_n$ be i.i.d. observations from repeated trials of the experiment. Let

$$N_j = \sum_{i=1}^{n} \mathbf{1}\left(X_i = j\right)$$

be the counts of the number each outcome is observed in the data.

Dimension of the Parameter

0/1 point (graded)
What must be true about d, the dimension of the model?

- $d = 1$
- $d = k - 1$
- $d \leq k - 1$ ✔
- (selected) $d \geq k - 1$

✖

Solution:

Since the asymptotic distribution of the MLE is non-degenerate, we must have $d \leq k$.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Degrees of Freedom

1/1 point (graded)
Consider the following χ^2 test statistics:

$$T_n = n \sum_{i=1}^{k} \frac{\left(N_i/n - P_\theta\left(i\right)\right)^2}{P_\theta\left(i\right)}.$$

Under the null hypothesis

$$H_0 : \theta = \theta^0$$

what is the number of degrees of freedom of the asymptotic distribution, which is a χ^2 distribution, of T_n?

Degrees of freedom: k-1 ✔ **Answer:** k-1

$k - 1$

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

Alternate Chi Square Test

1/1 point (graded)

Now, consider the alternate χ^2 test statistic below:

$$\widetilde{T_n} = n \sum_{i=1}^{k} \frac{\left(N_i/n - P_{\hat{\theta}}(i)\right)^2}{P_{\hat{\theta}}(i)}.$$

(The difference between T_n and $\widetilde{T_n}$ is that θ is changed to $\hat{\theta}$ as the parameter for P.)

Suppose that q_α is chosen such that

$$\lim_{n\to\infty} P_{\theta^0}\left(T_n > q_\alpha\right) = \alpha$$

where $T_n = n\sum_{i=1}^{k} \frac{\left(N_i/n - P_\theta(i)\right)^2}{P_\theta(i)}$ was defined as in the previous problem.

Which of the following must be true?

- $\lim_{n\to\infty} P_{\theta^0}\left(\widetilde{T_n} > q_\alpha\right) = \alpha$
- $\lim_{n\to\infty} P_{\theta^0}\left(\widetilde{T_n} > q_\alpha\right) < \alpha$
- $\lim_{n\to\infty} P_{\theta^0}\left(\widetilde{T_n} > q_\alpha\right) > \alpha$
- None of the above

Solution:

- Under the null hypothesis, $T_n^{(j=0)} \xrightarrow[n\to\infty]{(d)} \chi^2_{k-1}$ and coincides with the Wald's test of the multinomial statistical model. This follows from the discussion of Lecture 15.
- By estimating the parameter θ with MLE in the test statistic, we obtain an asymptotic χ^2 distribution with **fewer** degrees of freedom:

$$T_n^{(j)} \xrightarrow[n\to\infty]{(d)} \chi^2_{k-1-j}.$$

Therefore, the threshold $q_\alpha^{(j)}$ of the test $\psi_j\left(q_\alpha^{(j)}\right) = \mathbf{1}\left(T_n^{(j)} > q_\alpha^{(j)}\right)$ with asymptotic level α **decreases** with the number j of parameters that are estimated from data. Consequently, $\psi_{j+1}\left(q\right)$ is more **conservative** (rejects less often) than $\psi_j\left(q\right)$.

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Error and Bug Reports/Technical Issues

Hide Discussion

Topic: Midterm Exam 2:Midterm Exam 2 / 6.

Add a Post

Show all posts ▼ by recent activity ▼

[Polling] What are your answers for this page?
1) d = k-1 2) k-1 3) 1st choice
15

[Staff] - submission disappeared?
Dear Staff, my submission to the degrees of freedom question seems to be missing. I can't remember that I haven't submitted my answer (8). Could you please check in the s...
1