Physics
Carnegie Mellon

Note for Cosmostat project

Fromenteau Sebastien

April 19, 2016

1 χ^2 estimator and distribution

1.1 χ^2 estimator

1.2 χ^2 distribution

We will calculate the χ^2 probability distribution function (pdf) by induction over the number of dof. We will start to estimate the explicitly the pdf for $N_{dof} = 1$

1.2.1 $N_{dof} = 1$

For the case of $N_{dof} = 1$, we have $Y = x^2$ with $x \sim \mathcal{N}(0,1)$. We are interested in the probability distribution of the estimator Y.

$$
\begin{aligned}
P_1(Y \leq s) &= \int_{-\sqrt{s}}^{\sqrt{s}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx && (1)\\
P_1(Y \leq s) &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - \int_{-\infty}^{-\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx \right) && (2)\\
\int_{-\infty}^{-\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx &= 1 - \int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx && (3)\\
P_1(Y \leq s) &= \frac{1}{\sqrt{2\pi}} \left(2 \times \int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - 1 \right) && (4)\\
\int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx &= \frac{1}{2} - \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx && (5)\\
P_1(Y \leq s) &= \frac{1}{\sqrt{2\pi}} \left(2 \times \frac{1}{2} + 2 \times \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - 1 \right) && (6)\\
P_1(Y \leq s) &= \sqrt{\frac{2}{\pi}} \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx && (7)
\end{aligned}
$$

We want to calculate the probability the probability distribution function (pdf), so $p_1(Y = s)$ which is the derivative of the cumulative function $P_1(Y \leq s)$. So we have

$$
p_1(Y = s) = \frac{dP_1(Y \leq s)}{ds} \tag{8}
$$

We first have to do a substitution $z = x^2$ in order to have the expression in term of x^2 (which is the estimator : $Y = x^2$) and not in function of x. So we have

$$
z = x^2 \quad \Rightarrow \quad \frac{dz}{dx} = 2x \quad \Rightarrow \quad dx = \frac{dz}{2x} = \frac{dz}{2\sqrt{z}} \tag{9}
$$

The uper limit become $(\sqrt{s})^2 = s$

$$P_1(Y \le s) = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{\frac{2}{\pi}} \int_0^{s} \exp\left(-\frac{z}{2}\right) \frac{dz}{2\sqrt{z}} \tag{10}$$

$$p_1(Y = s) = \frac{dP_1(Y \le s)}{ds} = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{s}{2}\right) \frac{1}{2\sqrt{s}} \tag{11}$$

$$\boxed{p_1(Y = s) = \frac{1}{\sqrt{2\pi}\sqrt{s}} \exp\left(-\frac{s}{2}\right)} \tag{12}$$

This result will be fundamental for the rest of the demonstration bacause we will use it for each step of the induction.

1.2.2 $N_{dof} = 2$

$$p_2(Y = s) = \int_0^s p_1(s-\alpha) p_1(\alpha) d\alpha \tag{13}$$

$$p_2(Y = s) = \int_0^s \frac{1}{\sqrt{2\pi}\sqrt{s-\alpha}} \exp\left(-\frac{s-\alpha}{2}\right) \frac{1}{\sqrt{2\pi}\sqrt{\alpha}} \exp\left(-\frac{\alpha}{2}\right) d\alpha \tag{14}$$

$$p_2(Y = s) = \frac{1}{2\pi} \int_0^s \frac{1}{\sqrt{\alpha}\sqrt{s-\alpha}} \exp\left(-\frac{s-\alpha+\alpha}{2}\right) d\alpha \tag{15}$$

$$p_2(Y = s) = \frac{1}{2\pi} \exp\left(-\frac{s}{2}\right) \underbrace{\int_0^s \frac{1}{\sqrt{\alpha}\sqrt{s-\alpha}} d\alpha}_{=\pi} \tag{16}$$

$$\boxed{p_2(Y = s) = \frac{1}{2} \exp\left(-\frac{s}{2}\right)} \tag{17}$$

1.2.3 $N_{dof} = 3$

$$p_3(Y = s) = \int_0^s p_2(s-\alpha) p_1(\alpha) d\alpha \tag{18}$$

$$p_3(Y = s) = \int_0^s \frac{1}{2} \exp\left(-\frac{s-\alpha}{2}\right) \frac{1}{\sqrt{2\pi}\sqrt{\alpha}} \exp\left(-\frac{\alpha}{2}\right) d\alpha \tag{19}$$

$$p_3(Y = s) = \frac{1}{2\sqrt{2\pi}} \int_0^s \frac{1}{\sqrt{\alpha}} \exp\left(-\frac{s-\alpha+\alpha}{2}\right) d\alpha \tag{20}$$

$$p_3(Y = s) = \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{s}{2}\right) \underbrace{\int_0^s \frac{1}{\sqrt{\alpha}} d\alpha}_{=2\sqrt{s}} \tag{21}$$

$$\boxed{p_3(Y = s) = \frac{\sqrt{s}}{\sqrt{2\pi}} \exp\left(-\frac{s}{2}\right)} \tag{22}$$

1.2.4 $N_{dof} = 4$

$$p_4(Y = s) = \int_0^s p_2(s-\alpha) p_2(\alpha) d\alpha \tag{23}$$

$$p_4(Y = s) = \int_0^s \frac{1}{2} \exp\left(-\frac{s-\alpha}{2}\right) \frac{1}{2} \exp\left(-\frac{\alpha}{2}\right) d\alpha \tag{24}$$

$$p_4(Y = s) = \frac{1}{4} \exp\left(-\frac{s}{2}\right) \underbrace{\int_0^s d\alpha}_{=s} \tag{25}$$

$$\boxed{p_4(Y = s) = \frac{s}{4} \exp\left(-\frac{s}{2}\right)} \tag{26}$$

1.2.5 $N_{dof} = 5$

$$p_5(Y = s) = \int_0^s p_2(s-\alpha)p_3(\alpha)d\alpha \tag{27}$$

$$p_5(Y = s) = \int_0^s \frac{1}{2}\exp\left(-\frac{s-\alpha}{2}\right)\frac{\sqrt{\alpha}}{\sqrt{2\pi}}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{28}$$

$$p_5(Y = s) = \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{s}{2}\right)\underbrace{\int_0^s \sqrt{\alpha}d\alpha}_{=\frac{2}{3}s^{3/2}} \tag{29}$$

$$\boxed{p_5(Y = s) = \frac{s^{3/2}}{3\sqrt{2\pi}}\exp\left(-\frac{s}{2}\right)} \tag{30}$$

1.2.6 $N_{dof} = k$

We can see a pattern in the previous results which is:

$$p_k(Y = s) \propto s^{k/2-1}\exp\left(-\frac{s}{2}\right) \tag{31}$$

We have to use the induction to prove it. So we can start assuming that this relation is true for $N_{dof} = k$ and show that it's true for $N_{dof} = k+1$. I propose to use an even number for k. We can write the probability ti have $Y = s$ considering all the possibilities to have $Y_k = s - \alpha$ and $Y_1 = \alpha$.

$$p_{k+1}(Y = s) = \int_0^s p_k(s-\alpha)p_1(\alpha)d\alpha \tag{32}$$

$$p_{k+1}(Y = s) \propto \int_0^s (s-\alpha)^{k/2-1}\exp\left(-\frac{s-\alpha}{2}\right)\alpha^{-1/2}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{33}$$

$$p_{k+1}(Y = s) \propto \exp\left(-\frac{s}{2}\right)\int_0^s (s-\alpha)^{k/2-1}\alpha^{-1/2}d\alpha \tag{34}$$

We can develop $(s-\alpha)^{k/2-1}$ with the Binomial theorem as:

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}. \tag{35}$$

In order to simplify the notations we will use $\beta = k/2 - 1$ for the next development.

$$(s-\alpha)^\beta = \sum_{i=0}^{\beta}\binom{\beta}{i}s^i(-\alpha)^{\beta-i} \tag{36}$$

$$p_{k+1}(Y = s) \propto \exp\left(-\frac{s}{2}\right)\int_0^s \sum_{i=0}^{\beta}\binom{\beta}{i}s^i(-\alpha)^{\beta-i}\alpha^{-1/2}d\alpha \tag{37}$$

Since the sum converge (because the result is a binomial development for s and α which are finite numbers) we can reverse the sum and th integral. Moreover, the integral is over α so s^i can be remove from the integral too:

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right)\sum_{i=0}^{\beta}\binom{\beta}{i}s^i(-1)^{\beta-i}\int_0^s(\alpha)^{\beta-i}\alpha^{-1/2}d\alpha \tag{38}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right)\sum_{i=0}^{\beta}\binom{\beta}{i}s^i(-1)^{\beta-i}\int_0^s(\alpha)^{\beta-i-1/2}d\alpha \tag{39}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right)\sum_{i=0}^{\beta}\binom{\beta}{i}s^i(-1)^{\beta-i}\left[\frac{1}{\beta-i-1/2+1}\alpha^{\beta-i-1/2+1}\right]_0^s \tag{40}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right)\sum_{i=0}^{\beta}\binom{\beta}{i}(-1)^{\beta-i}\frac{1}{\beta-i-1/2+1}s^{\beta-i-1/2+1}\times s^i \tag{41}$$

$$p_{k+1}(Y=s) \propto \sum_{i=0}^{\beta}\binom{\beta}{i}(-1)^{\beta-i}\frac{1}{\beta-i-1/2+1}s^{\beta-1/2+1}\exp\left(-\frac{s}{2}\right) \tag{42}$$

We found finaly that the term in s doesn't depend anymore on i and so on the sum. The sum is just a factor. We rewrite with $\beta = k/2-1$ and we obtain that:

$$p_k(Y=s) \propto \left[\sum_{i=0}^{k/2-1}\binom{\beta}{i}(-1)^{\beta-i}\frac{1}{k/2-1-i-1/2+1}\right]\times s^{k/2-1-1/2+1}\exp\left(-\frac{s}{2}\right) \tag{43}$$

2 Fisher Matrix

The idea of the Fisher Matrix is to assume a fiducial cosmology (i.e a set of parameters assumed to be the good one) and see how well we are able to constrain these parameters using the second derivatien of the log-Likelihood.

In our context, we will use the relation between the Likelihood and the χ^2 estimator when the number of degree of freedom is large (when the χ^2 pdf is similar to a gaussian). In this case, we have $\mathcal{L} = \exp -\chi^2/2$

As we saw in the project, the general estimation for the χ^2 is:

$$\chi^2(\vec{\theta}|\vec{y}^D) = \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D\right)^t Cov^{-1}\left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D\right), \tag{44}$$

where $\vec{y}^D$ is the vector of data, $\vec{y}^{mod}(\vec{\theta})$ is the vector of points from the model for the parameters $\vec{\theta}$

3 Exercise

3.1 χ^2 distribution

1. Show that you retreive the same result for the pdf p_4Y using $p_4(Y=s) = \int_0^s p_2(s-\alpha)p_2(\alpha)d\alpha$
2. Estimate using a python code the value of s corresponding to 68% of the cumulative distribution function for $N_{dof} = 7$.
3. Knowing that the mean of the χ^2 pdf is $\mu = k$ and the variance is $\sigma^2 = 2k$, express the pdf for a gaussian following the same two first moment.
4. calculate the value for which the cumulative distribution function