Draft version June 21, 2020
Typeset using LaTeX **modern** style in AASTeX63

Notes With Henry

Will M. Farr[1] and Henry Shi[1]

[1] *Stony Brook University, Stony Brook NY 11794, United States*

1. RANDOM WALKS

Will posed the following problem to Henry. Consider a time series of values, x_i, over timesteps $i = 0, \ldots$. Let $x_0 = 0$, and

$$x_i = x_{i-1} + e_i \tag{1}$$

where e_i is a random variable with

$$e_i = \begin{cases} -1 & P = 1/2 \\ 1 & P = 1/2 \end{cases}. \tag{2}$$

Show that

$$\left\langle x_i \right\rangle = 0, \tag{3}$$

and find

$$\operatorname{Var} x_i. \tag{4}$$

Solution—First, the expectation value. Since expectations are linear, we have

$$\left\langle x_i \right\rangle = \left\langle x_{i-1} \right\rangle + \left\langle e_i \right\rangle = \left\langle x_{i-1} \right\rangle. \tag{5}$$

Since $x_0 = 0$, it follows (trivially) that $\left\langle x_0 \right\rangle = 0$, and therefore $\left\langle x_i \right\rangle = 0$ for all i.

Because the expectation is zero, we have

$$\operatorname{Var} x_i = \left\langle x_i^2 \right\rangle = \left\langle \left(x_{i-1} + e_i\right)^2 \right\rangle = \left\langle x_{i-1}^2 \right\rangle + 2 \left\langle x_{i-1} e_i \right\rangle + \left\langle e_i^2 \right\rangle. \tag{6}$$

Take all three terms in succession. First

$$\left\langle x_{i-1}^2 \right\rangle = \operatorname{Var} x_{i-1}. \tag{7}$$

will.farr@stonybrook.edu
henry.shi@stonybrook.edu

Next

$$\langle x_{i-1} e_i \rangle = \sum_{e_i = \pm 1} e_i x_{i-1} P(e_i) = \frac{1}{2}(x_{i-1} + (-x_{i-1})) = 0. \tag{8}$$

Finally

$$\langle e_i^2 \rangle = 1. \tag{9}$$

Putting it all together

$$\operatorname{Var} x_i = \operatorname{Var} x_{i-1} + 1. \tag{10}$$

Because $\operatorname{Var} x_0 = 0$, we have

$$\operatorname{Var} x_i = i. \tag{11}$$

2. χ^2

χ^2 is defined as

$$\chi^2 = \sum_i \left(\frac{d_i - f_i(\theta)}{\sigma_i} \right)^2 \tag{12}$$

where d_i are the measurements (the “data points”), f_i is the prediction for each measurement from the model, which depends on parameters θ, and σ_i is the uncertainty in each measurement. If we willing to assume that the “noise” is Gaussian, so that

$$d_i = f_i + n_i, \tag{13}$$

with the noise n_i distributed “normally”

$$n_i \sim N(0, \sigma_i), \tag{14}$$

then by the rules of probability, we have

$$p(\{d_i \mid i = 1, \ldots\} \mid \theta) = \frac{1}{\prod_i \sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2}\chi^2 \right]. \tag{15}$$

Hint(s): Since the only random thing here is the noise, remember that

$$p(d_i \mid \theta) = p(n_i = d_i - f_i(\theta)), \tag{16}$$

and that the Gaussian probability distribution for n_i is

$$p(n_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2} \left(\frac{n_i}{\sigma_i} \right)^2 \right]. \tag{17}$$

Show that the above formula in terms of χ^2 is the correct distribution for the data.

Bayes theorem:

$$p(d, \theta) = p(d \mid \theta)\, p(\theta) = p(\theta \mid d)\, p(d). \tag{18}$$

“Flipping the bar:”

$$p(\theta \mid d) = \frac{p(d \mid \theta)\, p(\theta)}{p(d)} \tag{19}$$

Your data consists of distance moduli

$$dm = m - M \tag{20}$$

where

$$dm = 5 \log_{10} \frac{d}{10\,\mathrm{pc}} \tag{21}$$

with

$$d = d_H \times \mathrm{stuff} = \frac{c}{H_0} \times \mathrm{stuff}, \tag{22}$$

so

$$dm = \ldots + \log_{10} H_0 + \ldots \tag{23}$$