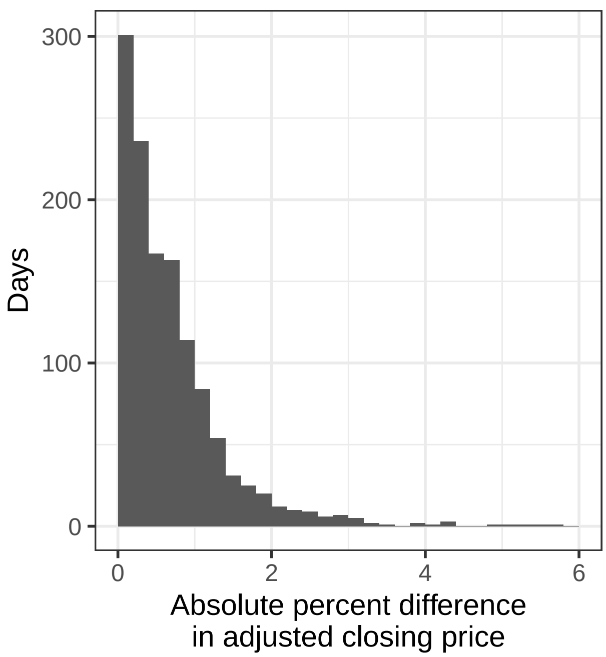
Announcements

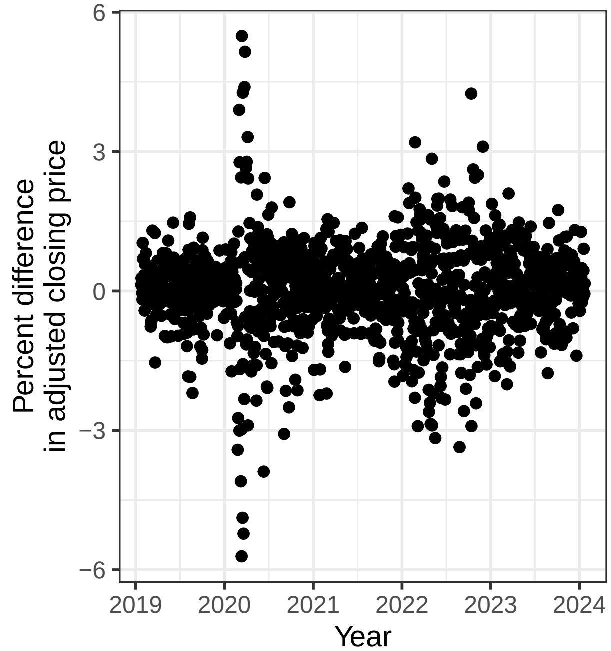
Make sure to sign in on the google form.

Pset 3 due 2/16.

Stonks

The following questions deal with the past 5 years of S&P 500 adjusted closing prices available here.





In this section, we will be modeling the day-to-day absolute percent differences in the adjusted closing price of the S&P 500 as $Y_1, ..., Y_n \sim \text{Expo}(\lambda)$.

1. Find the score of λ (the score is $\frac{\partial}{\partial \lambda}\ell(\lambda; \vec{y})$) in terms of the sample mean and verify that $E(s(\lambda^*; \vec{Y})) = 0$. (This equality was the last part of Neil's Thursday lecture.)

2. Show that the MLE of $\hat{\lambda}$ is consistent for λ. That is, show that $\hat{\lambda} \to \lambda$ as $n \to \infty$ by showing the MSE goes to 0, a LLN holds, making a claim using the CMT, or showing convergence directly.

3. In his book *The Black Swan*, Nassim Taleb argues that part of the reason for the 2008 financial crisis was a failure to model market fluctuations and assign sufficient probability to extreme events. Let us consider daily absolute differences above τ to be extreme events. Let $X_i = I(Y_i > \tau)$. Show that $\bar{X}$ is consistent for $p = P(Y_i > \tau)$ first by using MSE and then by using the law of large numbers.

4. Find the asymptotic distribution of $\bar{X}$ and its approximate distribution for large n in terms of λ.

5. Now, suppose we estimate $P(Y_i > \tau)$ with $\hat{p} = e^{-\hat{\lambda}\tau}$. Given that

$$\sqrt{n}(\hat{\lambda} - \lambda^*) \to \mathcal{N}\left(0, \lambda^{*2}\right)$$

(we'll see how to find this in future weeks), find the asymptotic distribution of $\hat{p}$ and its approximate distribution for large n.

6. The distributions in 4 and 5 should have the same mean. However, the variances are different. The following are the estimated standard errors of each estimator with $\tau = 5$. Explain why the results are what they are.

```
##  Xbar SE  phat SE
## 0.000759 0.000004
```

7. Though $\hat{p}$ is more efficient, it might be less robust. The following shows the MSEs of the two estimators for estimating $P(Y > \tau)$ when $Y \sim \text{Expo}(0.5)$ (the correct model) and $Y \sim \text{Log-Normal}(0.1, 1)$ (an incorrect model). Again, we have used $\tau = 5$ and $n = 100$ with 10^5 simulations.

```
##               Xbar    phat
## Log Normal 0.00061 0.00054
## Expo       0.00075 0.00041
```

Ty Mup

Ty Mup is taking an exam with n equally hard questions. He has a probability p_2 of getting each question right independently. However, there is also a $0 < p_1 < 1$ probability he sleeps through his alarm and misses the exam entirely. Let Y be the number of questions he gets right on his exam. (This distribution is called the zero-inflated binomial.)

1. Find $E(Y|Y > 0)$.

2. Unfortunately, the day is February 2nd in Punxsutawney and Ty is destined to repeat this day d times, scoring i.i.d Y_i on the exams. Find the likelihood function, the log-likelihood function, and the score for p_1 and p_2. (Hint: Let m be the number of 0s.)

3. Find a two-dimensional sufficient statistic (a two dimensional statistic that contains all the information about the likelihood).

4. Let B be the event that Ty sleeps through the alarm at least once. Show that as $d \to \infty$,

$$I_B \frac{d^{1/2}(\bar{Y} - (1-p_1)np_2)}{\sqrt{np_2(1-p_2)(1-p_1) + (np_2)^2 p_1(1-p_1)}} \xrightarrow{d} \mathcal{N}(0,1)$$