1. (Poisson Regression) The independent random variables $Y_i, i = 1, 2, ..n$,represent the outcomes of a Poisson experiment where the mean μ_i is propotional to the value of x_i. That is, $Y_i \sim Poisson(\mu_i)$ and $\mu_i = \gamma x_i$, Assume that the x_i, values are known constants.

a) Find the MLE of γ

$$L(\gamma) = \prod_{i=1}^{n} (\frac{\mu_i^{y_i}}{y_i!} e^{-\mu_i}) = \prod_{i=1}^{n} \frac{(\gamma x_i)^{y_i} e^{-\gamma x_i}}{y_i!} = \frac{\gamma^{\sum_{i=1}^n y_i} \prod_{i=1}^n x_i^{y_i}}{\prod_{i=1}^n y_i!} e^{-\gamma \sum_{i=1}^n x_i}, \quad y_i \in 0, 1, 2..$$

$$l(\gamma) = \ln \gamma \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} x_i^{y_i} - \sum_{i=1}^{n} \ln y_i! - \gamma \sum_{i=1}^{n} x_i$$

$$l'(\gamma) = \frac{\sum_{i=1}^n y_i}{\gamma} - \sum_{i=1}^{n} x_i \stackrel{\text{set}}{=} 0$$

$$\hat{\gamma}_{MLE} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$$

---

b) Find the mean and variance of $\hat{\gamma}_{MLE}$

For x_i are known constants. $Y_i \sim Poisson(\mu_i)$, $E[y_i] = Var[y_i] = \mu_i = \gamma x_i$,

$$E[\hat{\gamma}_{MLE}] = \frac{E[\sum_{i=1}^n y_i]}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n E[y_i]}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n \gamma x_i}{\sum_{i=1}^n x_i} = \gamma$$

For Y_i are independent random variables, $Cov(y_i, y_j) = 0, i \neq j$, $Var[\sum_{i=1}^n y_i] = \sum_{i=1}^n Var[y_i]$

$$Var[\hat{\gamma}_{MLE}] = \frac{Var[\sum_{i=1}^n y_i]}{(\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n Var[y_i]}{(\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n \gamma x_i}{(\sum_{i=1}^n x_i)^2} = \frac{\gamma}{\sum_{i=1}^n x_i}$$

---

2. Consider the regression model $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, .., n$. Find the maximum likelihood estimates of the paramiters if:

a)$\varepsilon_i \sim N(0, \sigma^2 x_i^2)$, independent for $i = 1, .., n$.

For $E[\varepsilon_i] = 0$, $Var[\varepsilon_i] = \sigma^2 x_i^2$, x_i and ε_i are independent,

$E[y_i] = E[\beta_0 + \beta_1 x_i] + E[\varepsilon_i] = \beta_0 + \beta_1 x_i$

$Var[y_i] = Var[\beta_0 + \beta_1 x_i] + Var[\varepsilon_i] = \sigma^2 x_i^2$

$Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2 x_i^2)$

$f_Y(y_i) = \frac{1}{x_i \sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2 x_i^2}(y_i - \beta_0 - \beta_1 x_i)^2}$

$$L(\beta_0, \beta_1, \sigma^2) = \prod_{i=1}^{n} (x_i \sqrt{2\pi\sigma^2})^{-1} e^{\sum_{i=1}^n \frac{-1}{\sigma^2 x_i^2}(y_i - \beta_0 - \beta_1 x_i)^2} = (2\pi\sigma^2)^{-\frac{n}{2}} (\prod_{i=1}^{n} x_i)^{-1} e^{\sum_{i=1}^n \frac{-1}{2\sigma^2 x_i^2}(y_i - \beta_0 - \beta_1 x_i)^2}$$

$$l(\beta_0,\beta_1,\sigma^2) = -\frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln(2\pi) - \ln(\prod_{i=1}^{n} x_i) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\frac{y_i}{x_i} - \frac{\beta_0}{x_i} - \beta_1)^2$$

Let $u_i = \frac{1}{x_i}, v_i = \frac{y_i}{x_i}$

$\frac{\partial l}{\partial \beta_1} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(v_i - u_i\beta_0 - \beta_1)(-1) \overset{\text{set}}{=} 0$

$$n\hat{\beta}_1 = \sum_{i=1}^{n} v_i - \hat{\beta}_0\sum_{i=1}^{n} u_i \implies \hat{\beta}_1 = \bar{v} - \bar{u}\hat{\beta}_0 \quad (1)$$

$\frac{\partial l}{\partial \beta_0} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(v_i - u_i\beta_0 - \beta_1)(-u_i) \overset{\text{set}}{=} 0$

$$\hat{\beta}_1\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} u_i v_i - \hat{\beta}_0\sum_{i=1}^{n} u_i^2 \implies n\bar{u}\hat{\beta}_1 = \sum_{i=1}^{n} u_i v_i - \hat{\beta}_0\sum_{i=1}^{n} u_i^2 \quad (2)$$

The solution of (1) and (2) is

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} u_i v_i - n\bar{u}\bar{v}}{\sum_{i=1}^{n} u_i^2 - n\bar{u}^2} = \frac{S_{uv}}{S_{uu}}$$

$$\hat{\beta}_1 = \bar{v} - \bar{u}\frac{S_{uv}}{S_{uu}}$$

$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(v_i - u_i\beta_0 - \beta_1)^2 \overset{\text{set}}{=} 0$

$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(v_i - u_i\beta_0 - \beta_1)^2 = \frac{1}{n}\sum_{i=1}^{n}(v_i - u_i\beta_0 - \bar{v} + \bar{u}\hat{\beta}_0)^2$

$= \frac{1}{n}\left[\sum_{i=1}^{n}(v_i - \bar{v})^2 + \beta_0^2\sum_{i=1}^{n}(u_i - \bar{u})^2 - 2\beta_0\sum_{i=1}^{n}(u_i - \bar{u})(v_i - \bar{v})\right]$

$$= \frac{1}{n}\left[S_{vv} + (\frac{S_{uv}}{S_{uu}})^2 S_{uu} - 2\frac{S_{uv}}{S_{uu}}S_{uv}\right] = \frac{1}{n}\left[S_{vv} - \frac{S_{uv}^2}{S_{uu}}\right]$$

For $u_i = \frac{1}{x_i}, v_i = \frac{y_i}{x_i}$

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} u_i v_i - n\bar{u}\bar{v}}{\sum_{i=1}^{n} u_i^2 - n\bar{u}^2} = \frac{\sum_{i=1}^{n}(y_i/x_i^2) - n\overline{(1/x)}\,\overline{(y/x)}}{\sum_{i=1}^{n}(1/x_i)^2 - n\overline{(1/x)}^2}$$

$$\hat{\beta}_1 = \frac{\bar{v}\sum_{i=1}^{n} u_i^2 - \bar{u}\sum_{i=1}^{n} u_i v_i}{\sum_{i=1}^{n} u_i^2 - n\bar{u}^2} = \frac{\overline{(y/x)}\sum_{i=1}^{n}(1/x_i)^2 - \overline{(1/x)}\sum_{i=1}^{n}(y_i/x_i^2)}{\sum_{i=1}^{n}(1/x_i)^2 - n\overline{(1/x)}^2}$$

$$\hat{\sigma}^2 = \frac{1}{n}\left\{\sum_{i=1}^{n}(y_i^2/x_i^2) - n\overline{(y/x)}^2 - \frac{\left[\sum_{i=1}^{n}(y_i/x_i^2) - n\overline{(1/x)}\,\overline{(y/x)}\right]^2}{\sum_{i=1}^{n}(1/x_i)^2 - n\overline{(1/x)}^2}\right\}$$

---

b) $\varepsilon_i \sim i.i.d.\ f(\varepsilon;\lambda) = \frac{\lambda}{2}e^{-\lambda|\varepsilon|}$.

$\varepsilon_i \sim Laplace(0, \lambda)$, $E[\varepsilon_i] = 0$

$\varepsilon_i = y_i - \beta_0 - \beta_1 x_i$

$E[y_i] = E[\beta_0 + \beta_1 x_i] + E[\varepsilon_i] = \beta_0 + \beta_1 x_i$

$Y_i \sim Laplace(\beta_0 + \beta_1 x_i, \lambda)$

$f_Y(y_i) = \frac{\lambda}{2} e^{-\lambda |y_i - \beta_0 - \beta_1 x_i|}$

$$L(\beta_0, \beta_1, \lambda) = \prod_{i=1}^{n} (\frac{\lambda}{2} e^{-\lambda |y_i - \beta_0 - \beta_1 x_i|}) = \lambda^n 2^{-n} e^{-\lambda \sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_i|}$$

$$l(\beta_0, \beta_1, \lambda) = n \ln \lambda - n \ln 2 - \lambda \sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_i|$$

$\frac{\partial l}{\partial \beta_0} = -\lambda \sum_{i=1}^{n} \begin{cases} -1 & \text{if } y_i > \beta_0 + \beta_1 x_i \\ 0 & \text{if } y_i = \beta_0 + \beta_1 x_i \\ 1 & \text{if } y_i < \beta_0 + \beta_1 x_i \end{cases} = \lambda \sum_{i=1}^{n} sgn(y_i - \beta_0 - \beta_1 x_i)$

$\frac{\partial l}{\partial \beta_1} = -\lambda \sum_{i=1}^{n} \begin{cases} -x_i & \text{if } y_i > \beta_0 + \beta_1 x_i \\ 0 & \text{if } y_i = \beta_0 + \beta_1 x_i \\ x_i & \text{if } y_i < \beta_0 + \beta_1 x_i \end{cases} = \lambda \sum_{i=1}^{n} sgn(y_i - \beta_0 - \beta_1 x_i) x_i$

To minimize the total absolute deviations,

$\beta_0 + \beta_1 x_i = y_m$ (median).

$1/\lambda$ is the Mean Absolute Deviation from the median.

Assume $\varepsilon_1, .., \varepsilon_n$ are ordered. Let $\varepsilon_1, .., \varepsilon_j < 0$, $\varepsilon_{j+1}, .., \varepsilon_n > 0$,

$\sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_i| = -\sum_{i=1}^{j} (y_i - \beta_0 - \beta_1 x_i) + \sum_{i=j+1}^{n} (y_i - \beta_0 - \beta_1 x_i)$

$\frac{\partial l}{\partial \beta_0} = -j\lambda + (n-j)\lambda = (n-2j)\lambda \overset{\text{set}}{=} 0 \implies j = \frac{n}{2},\ x_j = x_m \quad (median)$

$\frac{\partial l}{\partial \beta_1} = -\lambda \sum_{i=1}^{j} x_i + \lambda \sum_{i=j+1}^{n} x_i = \lambda(\sum_{i=j+1}^{n} x_i - \sum_{i=1}^{j} x_i) \overset{\text{set}}{=} 0 \implies x_j = \bar{x}$

$$\begin{cases} y_m - \hat{\beta}_0 - \hat{\beta}_1 x_m = 0 \\ \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0 \end{cases} \implies \begin{cases} \hat{\beta}_0 = \frac{x_m \bar{y} - y_m \bar{x}}{x_m - \bar{x}} \\ \hat{\beta}_1 = \frac{y_m - \bar{y}}{x_m - \bar{x}} \end{cases}$$

$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_i| \overset{\text{set}}{=} 0$

$$\hat{\lambda}_{MLE} = \frac{n}{\sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_i|} = \frac{n}{\sum_{i=1}^{n} |y_i - y_m|} = \frac{n}{\sum_{i=1}^{n} |y_i - \frac{x_m \bar{y} - y_m \bar{x}}{x_m - \bar{x}} - \frac{y_m - \bar{y}}{x_m - \bar{x}} x_i|}$$

---

3. Finde the finite breakdown point and the infinite breakdown point for

a) the Mean Absolute Deviation, or $\frac{1}{n} \sum_{i=1}^{n} |X_i - \bar{X}|$.

The finite breakdown point is the smallest proportion m/n of the sample values such that $|\hat{\theta}^* - \hat{\theta}|$ can be made arbitarily large by corrupting m data values and computing $\hat{\theta}^*$, where n is the samle size, $\hat{\theta}$ is the estimator. The limit as $n \to \infty$ is called the breakdown point.

Assume the $X_1, .., X_n$ are ordered. Let $X_j < \bar{X},\ X_{j+1} > \bar{X}$.

Replace X_n with a arbitrarily large X_n^*.

If $X_n^* > nX_n - \sum_{i=1}^{n-1} X_i$, then $\bar{X}^* > X_n$

$|\hat{\theta}^* - \hat{\theta}| = |\frac{1}{n}\sum_{i=1}^{n-1}(-X_i + \bar{X}^*) + \frac{1}{n}(X_n^* - \bar{X}^*) - \frac{1}{n}\sum_{i=1}^{j}(-X_i + \bar{X}) - \frac{1}{n}\sum_{i=j+1}^{n-1}(X_i - \bar{X}) - \frac{1}{n}(X_n - \bar{X})|$

$= \frac{1}{n}|\sum_{i=1}^{j}(\bar{X}^* - \bar{X}) + \sum_{i=j+1}^{n-1}(\bar{X}^* + \bar{X}) - 2\sum_{i=j+1}^{n-1} X_i - (\bar{X}^* - \bar{X}) + X_n^* - X_n|$

$= \frac{1}{n}|(n-2)\bar{X}^* + (n-2j)\bar{X} - 2\sum_{i=j+1}^{n-1} X_i + X_n^* - X_n|$

$= \frac{1}{n^2}|(n-2)(\sum_{i=1}^{n-1} X_i + X_n^*) + (n-2j)(\sum_{i=1}^{n-1} X_i + X_n) - 2n\sum_{i=j+1}^{n-1} X_i + nX_n^* - nX_n|$

$= \frac{1}{n^2}|(2n-2j-2)\sum_{i=1}^{n-1} X_i + (2n-2)X_n^* - 2jX_n - 2n\sum_{i=j+1}^{n-1} X_i|$

$= \frac{2}{n^2}|(n-j-1)\sum_{i=1}^{n-1} X_i + (n-1)X_n^* - jX_n - n\sum_{i=j+1}^{n-1} X_i|$

$$= \frac{2}{n^2}|n\sum_{i=1}^{j} X_i - j\sum_{i=1}^{n} X_i - \sum_{i=1}^{n-1} X_i + (n-1)X_n^*|$$

We just need corrupt one value in order to corrupt MAD.

The finite breakdown point $= \frac{1}{n}$

The infinite breakdown point $= \lim_{n\to\infty}\frac{1}{n} = 0$

---

b) the Median Absolute Deviation, or Median$\{|X_1 - \bar{X}|, .., |X_n - \bar{X}|\}$.

Assume the $X_1, .., X_n$ are ordered. Let $X_j < \bar{X} < X_{j+1}$

$$\{|X_1 - \bar{X}|, .., |X_n - \bar{X}|\} = \{-X_1 + \bar{X}, .., -X_j + \bar{X}, X_{j+1} - \bar{X}, .., X_n - \bar{X}\}$$

Rearrange the order

$$\begin{cases} \{-X_j + \bar{X}, .., -X_1 + \bar{X}\} & (1) \\ \{X_{j+1} - \bar{X}, .., X_n - \bar{X}\} & (2) \end{cases}$$

$\hat{\theta}$ might be $-X_k + \bar{X} \in (-X_j + \bar{X}, .., -X_1 + \bar{X})$,

or $X_k - \bar{X} \in (X_{j+1} - \bar{X}, .., X_n - \bar{X})$.

k depend on both of the orders (1) and (2).

Replace X_n with a arbitrarily large X_n^*, $\bar{X}^* >> X_n$

$$\{|X_1 - \bar{X}^*|, .., |X_n^* - \bar{X}^*|\} = \{-X_1 + \bar{X}^*, .., -X_{n-1} + \bar{X}^*, X_n^* - \bar{X}^*\}$$

When n is even, $\hat{\theta}^*$ is $-X_{\frac{n}{2}} + \bar{X}^* > -X_1 + \bar{X}$ and $> X_n - \bar{X}$.

When n is odd, $\hat{\theta}^*$ is $-X_{\frac{n+1}{2}} + \bar{X}^* > -X_1 + \bar{X}$ and $> X_n - \bar{X}$.

Therefore, we just need corrupt one value in order to corrupt MAD.

The finite breakdown point $= \frac{1}{n}$

The infinite breakdown point $= \lim_{n\to\infty} \frac{1}{n} = 0$

---

4. Assume that $X_1, X_2, ..X_n$ are i.i.d. $Uniform(a, b)$. Find the asymptotic relative efficiency of the sample median to the sample mean.

For $X \sim Unif(a, b)$, $E[X] = \frac{a+b}{2}$, $Var[X] = \frac{(b-a)^2}{12}$, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$,

$E[\bar{X}] = E[\frac{1}{n}\sum_{i=1}^{n} x_i] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \frac{1}{n}\sum_{i=1}^{n} \frac{a+b}{2} = \frac{a+b}{2}$

For X_i are independent,

$Var[\bar{X}] = Var[\frac{1}{n}\sum_{i=1}^{n} x_i] = \frac{1}{n^2}\sum_{i=1}^{n} Var[x_i] = \frac{1}{n^2}\sum_{i=1}^{n} \frac{(b-a)^2}{12} = \frac{(b-a)^2}{12n}$

$$\bar{X} \sim N(\frac{a+b}{2}, \frac{(b-a)^2}{12n})$$

For large n, the sample median $m_n \approx N(M, \frac{1}{4nf^2(M)})$, where M is the population median.

$f(x) = \frac{1}{b-a}$ is the p.d.f. of X

$E[m_n] = M$

$Var[m_n] = \frac{1}{4nf^2(M)} = \frac{(b-a)^2}{4n}$

The asymptotic relative efficiency of m_n to $\bar{X}$

$$= \frac{Var[\bar{X}]}{Var[m_n]} = \frac{\frac{(b-a)^2}{12n}}{\frac{(b-a)^2}{4n}} = \frac{1}{3}$$

Therefore, the sample mean is asymptotic more efficiency than sample median.