8.2 The Chi Squared Distribution

Recall that the MLE estimator for the variance σ^2 of a Normal distribution with known mean μ is

$$\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \mu)^2$$

We now derive/define the sampling distribution for $\hat{\sigma}^2$.

Def: For each positive integer m, the gamma distribution with parameters $\alpha = m/2$ and $\beta = 1/2$ is called the χ^2 distribution with m degrees of freedom. As such, the density for $X \sim \chi^2(m)$ is

$$f(x) = \frac{1}{2^{m/2}\Gamma(m/2)} x^{m/2} e^{-x/2} \frac{1}{x} \qquad \text{for } x \geq 0.$$

And the mean and variance of X are

$$E[X] = m \qquad \mathrm{Var}(X) = 2m.$$

The MGF of X is

$$M_X(t) = \left(\frac{1}{1-2t}\right)^{m/2}$$

Ex 1: Investigate the density for several χ^2 distributions.

Thm: If $X_1, \ldots, X_k$ are independent, with $X_i \sim \chi^2(m_i)$, then $X_1 + \cdots + X_k$ has χ^2 distribution with $m = m_1 + \cdots + m_k$ degrees of freedom.

Proof. Use MGFs. □

What is the link between the Normal distribution and the χ^2 distribution?

Thm: Suppose $Z \sim N(0,1)$. Then $Y = Z^2$ is $\chi^2(1)$.

Proof. Recall the change of variables formula. If g is invertible and $Y = g(X)$, then

$$f_Y(y) = f_X(g^{-1}(y)) \left|\frac{dx}{dy}\right|.$$

If g is differentiable, but not invertible, then f_Y is given by a sum across the pre-images of y.

Here, we first consider $x \geq 0$. Let $y = x^2$, so that $\sqrt{y} = x$. Then

$$\begin{aligned} f_Y(y) =& f_X(g^{-1}(y)) \left|\frac{dx}{dy}\right| \\ =& \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left|\frac{1}{2\sqrt{y}}\right| \\ =& \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} \\ =& \frac{1}{2\sqrt{2\pi}} y^{1/2} e^{-y/2} \frac{1}{y} \end{aligned}$$

Now, if $x < 0$, then $-\sqrt{y} = x$. But observe that the preceding formula still holds in this case. □

Ex 2: True or False? If $X \sim \chi^2(1)$, then $\sqrt{X} \sim N(0,1)$?

Cor: If $X_1, \ldots, X_n$ are iid $N(0,1)$, then $X_1^2 + \cdots + X_n^2$ is χ^2 with n degrees of freedom.

Thm: Let $X_1, \dots, X_n$ be a sample from a Normal population with mean μ and variance σ^2. Define $Z_i = (X_i - \mu)/\sigma$. Then $\sum Z_i^2$ is χ^2 with n degrees of freedom, as is $n\hat{\sigma}^2/\sigma^2$.

Proof. Note that

$$\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{n}{\sigma^2}\frac{1}{n}\sum (X_i - \mu)^2 = \sum \left(\frac{X_i - \mu}{\sigma}\right)^2 = \sum Z_i^2$$

and that $Z_i \sim N(0,1)$ by location-scale transformations. The result follows by the previous corollary. □

Cor: The estimator $\hat{\sigma}^2$ is Gamma$(n/2, n/(2\sigma^2))$, which has mean and variance of

$$E[\hat{\sigma}^2] = \sigma^2 \qquad \text{Var}(\hat{\sigma}^2) = \frac{2\sigma^4}{n}$$

Note: $\hat{\sigma}^2$ is an *unbiased* estimator of σ^2 when the mean is known.

8.3 Joint Distribution of the Sample Mean and Sample Variance

Now, suppose we are in the more realistic scenario where both μ and σ^2 in $N(\mu, \sigma^2)$ are unknown parameters to be estimated. Recall that the MLE $\hat{\theta}$ for $\theta = (\mu, \sigma^2)$ is

$$\hat{\theta} = \left(\bar{X}, \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)$$

We now calculate the joint sampling distribution of $\hat{\theta}$.

Thm: If $X_1, \dots, X_n$ are a random sample from $N(\mu, \sigma^2)$, the estimators $\bar{X} = \frac{1}{n}\sum X_i$ and $\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$ are independent. The marginal distribution of $\bar{X}$ is $N(\mu, \sigma^2/n)$, and the marginal distribution of $\frac{n}{\sigma^2}\hat{\sigma}^2$ is $\chi^2(n-1)$.

To prove this theorem, we make us of the Multivariate Normal Distribution and the following essential properties:

Def: A random vector $\mathbf{X}$ is said to be multivariate Normal if every linear combination of coordinates is Normally distributed.

Thm: A MVN $\mathbf{X}$ is completely determined by its mean vector $\boldsymbol{\mu}$ and covariance matrix Σ:

$$\boldsymbol{\mu} = (E[X_1], E[X_2], \dots, E[X_n]) \qquad \Sigma = (\text{Cov}(X_i, X_j))$$

Cor: If $\mathbf{X}$ is MVN, then $X_1, \dots, X_n$ are independent of one another if and only $\text{Cov}(X_i, X_j) = 0$ for all $i \neq j$.

Proof. The vector $(\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})$ is multivariate normal, since every linear combination of its coordinates are a linear combination of $X_1, \dots, X_n$ (which are iid $N(\mu, \sigma^2)$ by assumption). Moreover, $E[X_i - \bar{X}] = 0$ by linearity. We now computed the covariances of $\bar{X}$ with $X_i - \bar{X}$. Using the bilinearity of covariance:

$$\text{Cov}(\bar{X}, X_i - \bar{X}) = \text{Cov}(\bar{X}, X_i) - \text{Cov}(\bar{X}, \bar{X})$$

For the first term,

$$\text{Cov}(\bar{X}, X_i) = \frac{1}{n}\text{Cov}\left(X_1 + \cdots + X_n, X_i\right) = \frac{1}{n}\text{Cov}(X_i, X_i) = \frac{1}{n}\text{Var}(X_i) = \frac{\sigma^2}{n}$$

For the second term,

$$\text{Cov}(\bar{X}, \bar{X}) = \text{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Hence,

$$\text{Cov}(\bar{X}, X_i - \bar{X}) = \text{Cov}(\bar{X}, X_i) - \text{Cov}(\bar{X}, \bar{X}) = 0$$

But this means that $\bar{X}$ is uncorrelated with every other component of $(X_1 - \bar{X}, \ldots, X_n - \bar{X})$. Since uncorrelated implies independent for Multivariate Normal vectors, then $\bar{X}$ is independent of $(X_1 - \bar{X}, \ldots, X_n - \bar{X})$. But $\hat{\sigma}^2$ is a function of this vector, and so $\bar{X}$ is independent of $\hat{\sigma}^2$ as well.

Now, assume that $X_1, \ldots, X_n$ are **standard** Normal (we'll use this to bootstrap up to the general case). Note that

$$\sum X_i^2 = \sum (X_i - \bar{X})^2 + n\bar{X}^2$$

We now compute the MGFs of both sides of the equation. Since the sample mean and variance are independent, then the MGF of their sum is the product of their MGFs. Moreover,

$$\sum X_i^2 \sim \chi^2(n) \qquad n\bar{X}^2 \sim \chi^2(1)$$

Thus, letting M be the MGF of $\sum (X_i - \bar{X})^2$,

$$\left(\frac{1}{1-2t}\right)^{n/2} = M(t)\left(\frac{1}{1-2t}\right)^{1/2}$$

and so

$$M(t) = \left(\frac{1}{1-2t}\right)^{(n-1)/2}$$

which is the MGF of $\chi^2(n-1)$.

For the general case with $X_i \sim N(\mu, \sigma^2)$, let $X_i = \mu + \sigma Z_i$ for $Z_i \sim N(0,1)$. Then

$$\sum (X_i - \bar{X})^2 = \sum \left(\mu + \sigma Z_i - (\mu + \sigma \bar{Z})\right)^2 = \sigma^2 \sum (Z_i - \bar{Z})^2$$

It follows that

$$\frac{n}{\sigma^2}\hat{\sigma}^2 = \frac{1}{\sigma^2}\sum (X_i - \bar{X})^2 = \frac{1}{\sigma^2}\sigma^2 \sum (Z_i - \bar{Z})^2 \sim \chi^2(n-1)$$

as desired.

Also, we note that this gives a quick proof that $E[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2$. $\square$