$X_1 \ldots X_n$ form a random sample $N(\mu, \sigma^2)$

$$\bar{X}_n \sim N(\mu, \sigma^2/n)$$

$$\frac{n\hat{\sigma}_0^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_{(n)} \quad , \quad \hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

$\bar{X}_n$

Theorem for joint distr. of $\bar{X}_n$ and $\hat{\sigma}^2$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

$X_1 \ldots X_n$ form random sample $N(\mu, \sigma^2)$

$\bar{X}_n$ and $\hat{\sigma}^2$ are independent random variables.

$$\left.\begin{aligned} &\bar{X}_n \sim N(\mu, \sigma^2/n) \\ &\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}_n}{\sigma}\right)^2 \sim \chi^2_{(n-1)} \end{aligned}\right\} \text{independent.}$$

$$\left.\begin{aligned} &\bar{X}_n \sim N(\mu, \sigma^2/n) \\ &\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)} \end{aligned}\right\} \text{independent.}$$

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

Example: lactic acid in cheese.

X_i : concentration of lactic acid in cheese.

$X_i \sim N(\mu, \sigma^2)$

a) does the sample variance, $\hat{\sigma}^2$, underestimate the variance of the r.v $\cancel{X}$, σ^2 ?

$$P(\hat{\sigma}^2 < \sigma^2) = ?$$

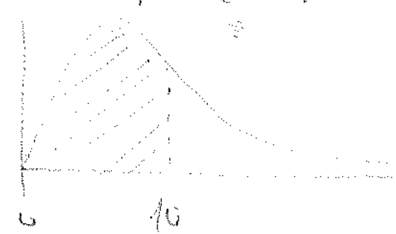
We know that $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)}$ (theorem of joint distr)

$$P(\hat{\sigma}^2 < \sigma^2) = P\left(\frac{\hat{\sigma}^2}{\sigma^2} < 1\right) = P\Big(\underbrace{\frac{n\hat{\sigma}^2}{\sigma^2}}_{\chi^2_{(n-1)}} < n\Big)$$

$$= P(Y < n) \quad , \quad Y \sim \chi^2_{(n-1)}$$

$$= P(Y < 10)$$

$$= 0.649$$



pchisq(10, df = 9)

$X_1, \dots X_n$ form a random sample from $N(\mu, \sigma^2)$

From theorem of joint distr. for $\bar{X}_n$, $\hat{\sigma}^2$.

$$\left.\begin{array}{l} \bar{X}_n \sim N(\mu, \sigma^2/n) \\ \dfrac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)} \end{array}\right\} \text{ are independent.}$$

$$\hat{\mu} = \bar{X}_n$$
$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

$$P\left(|\hat{\mu} - \mu| < \frac{\sigma}{5} \text{ and } |\hat{\sigma} - \sigma| < \frac{\sigma}{5}\right) \geq \frac{1}{2}.$$

$$P\left(|\bar{X}_n - \mu| < \frac{\sigma}{5} \text{ and } \left|\sqrt{\hat{\sigma}^2} - \sigma\right| < \frac{\sigma}{5}\right)$$

$$= P\left(|\bar{X}_n - \mu| < \frac{\sigma}{5}\right) P\left(\left|\sqrt{\hat{\sigma}^2} - \sigma\right| < \frac{\sigma}{5}\right)$$, because $\bar{X}_n$ and $\hat{\sigma}^2$ are independent.

$$= P\left(-\frac{\sigma}{5} < \bar{X}_n - \mu < \frac{\sigma}{5}\right) P\left(-\frac{\sigma}{5} < \sqrt{\hat{\sigma}^2} - \sigma < \frac{\sigma}{5}\right)$$

$$= P\left(-\frac{\sigma \cdot \sqrt{n}}{5 \cdot \sigma} < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} < \frac{\sigma\sqrt{n}}{5\sigma}\right) P\left(\frac{4\sigma}{5} < \sqrt{\hat{\sigma}^2} < \frac{6\sigma}{5}\right)$$

$$= P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) P\left(\frac{4}{5} < \frac{\sqrt{\hat{\sigma}^2}}{\sigma} < \frac{6}{5}\right), \quad Z \sim N(0,1).$$

$$= P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) P\left(\frac{16}{25} < \frac{\hat{\sigma}^2}{\sigma^2} < \frac{36}{5}\right)$$

$$= P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) P\left(\frac{16 \cdot n}{25} < \frac{n\hat{\sigma}^2}{\sigma^2} < \frac{36 \cdot n}{5}\right)$$

$$= P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) P\left(\frac{16 \cdot n}{25} < Y < \frac{36 \cdot n}{5}\right), \quad Z \sim N(0,1), \; Y \sim \chi^2_{(n-1)}$$

$$= \underbrace{P\left(-\frac{\sqrt{21}}{5} < Z < \frac{\sqrt{21}}{5}\right)}\,\underbrace{P\left(\frac{16 \cdot 21}{25} < Y < \frac{36 \cdot 21}{25}\right)}$$

$$= 0.6406 \cdot 0.7917 = 0.5070.$$

$X_1 .. X_n$ form random sample ~~$\sim$~~ $N(\mu, \sigma^2)$

$$\left.\begin{array}{l} \bar{X}_n \sim N(\mu, \sigma^2/n) \\ \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-1)} \end{array}\right\rangle \text{ independent}$$

$$\Rightarrow \left.\begin{array}{l} \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \\ \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1) \end{array}\right\rangle \text{ independent.}$$

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}} \sim ?$$

---

t- distribution.

$$X = \frac{Z}{\sqrt{\frac{Y}{m}}}, \quad \left.\begin{array}{l} Z \sim N(0,1) \\ Y \sim \chi^2(m) \end{array}\right\rangle \text{ are independent}$$

$$Z = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}, \quad Y = \frac{n\hat{\sigma}^2}{\sigma^2} = \frac{n \cdot \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2}{\sigma^2}$$

$$X = \frac{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}}{\sqrt{\left\{\frac{\frac{\sum(X_i - \bar{X}_n)^2}{\sigma^2}}{n-1}\right\}}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \cdot \frac{1}{\sqrt{\frac{1}{\sigma^2}}\sqrt{\frac{\sum(X_i - \bar{X}_n)^2}{\cancel{\sigma^2} n-1}}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\frac{\sum(X_i - \bar{X}_n)^2}{n-1}}}$$

$$= \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma'}$$ and it has the t distribution with $n-1$ degrees of freedom.