Show $\overline{X}$ and S^2 are independent

(Under the assumption the random sample is normally distributed)

A well known result in statistics is the independence of $\overline{X}$ and S^2 when $X_1, X_2, \cdots, X_n \sim \mathrm{N}\big(\mu, \sigma^2\big)$. This handout presents a proof of the result using a series of results. First, a few lemmas are presented which will allow succeeding results to follow more easily. In addition, the distribution of $\frac{(n-1)S^2}{\sigma^2}$ is derived.

Definition 1. *The sample variance is defined as*

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Lemma 1. *The sum of the squares of the random variables $X_1, X_2, \cdots, X_n$ is*

$$\sum_{i=1}^{n} X_i^2 = (n-1)S^2 + n\overline{X}^2$$

Proof. By Definition 1,

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 = \sum_{i=1}^{n} X_i^2 - 2\overline{X}\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\overline{X}^2 = \sum_{i=1}^{n} X_i^2 - 2n\overline{X}^2 + n\overline{X}^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$$

It follows that

$$\sum_{i=1}^{n} X_i^2 = (n-1)S^2 + n\overline{X}^2$$

□

Lemma 2. *The sum of squares of the random variables $X_1, X_2, \cdots X_n$ centered about the mean, μ, is*

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - \mu)^2$$

Proof. The sum of squares can be simplified as

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n} X_i^2 - 2\mu\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\mu^2$$

$$= \sum_{i=1}^{n} X_i^2 - 2n\mu\overline{X} + n\mu^2 \tag{1}$$

By Lemma 1, (1) simplifies to

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n} X_i^2 - 2n\mu\overline{X} + n\mu^2 = (n-1)S^2 + n\overline{X}^2 - 2n\mu\overline{X} + n\mu^2$$

$$= (n-1)S^2 + n(\overline{X} - \mu)^2 \tag{2}$$

$$= \sum_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - \mu)^2$$

□

Lemma 3. *If* $Z \sim \mathrm{N}(0,1)$*, then* $Z^2 \sim \chi^2(1)$.

Proof. The moment generating function for Z^2 is defined as

$$
\begin{aligned}
M_{Z^2}(t) = \mathrm{E}\Big(e^{tZ^2}\Big) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tz^2} e^{-z^2/2}\, dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(1-2t)z^2\right] dz \\
&= \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\frac{z^2}{\frac{1}{1-2t}}\right] dz}_{\text{kernel of a } N\left(0, \frac{1}{1-2t}\right)} \\
&= \frac{1}{\sqrt{1-2t}} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\left(\frac{1}{\sqrt{1-2t}}\right)} \exp\left[-\frac{1}{2}\frac{z^2}{\frac{1}{1-2t}}\right] dz}_{\text{integrates to 1}} \\
&= \frac{1}{\sqrt{1-2t}} \\
&= (1-2t)^{-1/2}
\end{aligned}
$$

Note that this is the moment generating function for a χ^2 random variable with one degree of freedom. Hence,

$$Z^2 \sim \chi^2(1)$$

□

Lemma 4. *Suppose* $X_1, X_2, \cdots, X_n$ *are independent and identically distributed* $\chi^2(1)$ *random variables. It follows that*

$$Y = \sum_{i=1}^{n} X_i \sim \chi^2(n)$$

Proof. The moment generating function of X_i is

$$M_{X_i}(t) = (1-2t)^{-1/2}.$$

It follows that the moment generating function for Y is

$$
\begin{aligned}
M_Y(t) = \mathrm{E}\big[e^{tY}\big] = \mathrm{E}\big[e^{tX_1+tX_2+\cdots+tX_n}\big] &= \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n}(1-2t)^{-1/2} \\
&= (1-2t)^{-\sum_{i=1}^{n} 1/2} \\
&= (1-2t)^{-n/2}
\end{aligned}
$$

It follows that this is the MGF for a χ^2 distribution with n degrees of freedom. Hence,

$$Y = \sum_{i=1}^{n} X_i \sim \chi^2(n)$$

□

Theorem 1. *Suppose* $X_1, X_2, \cdots, X_n$ *is a random sample from a normal distribution with mean,* μ*, and variance,* σ^2*. It follows that the sample mean,* $\overline{X}$*, is independent of* $X_i - \overline{X}$*,* $i = 1, 2, \cdots, n$*.*

Proof. The joint distribution of $X_1, X_2, \cdots, X_n$ is

$$f_X(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2\right\}$$

Transform the random variables X_i, $i = 1, 2, \cdots, n$ to

$$\begin{aligned} Y_1 &= \overline{X} & \overline{X} &= Y_1 \\ Y_2 &= X_2 - \overline{X} & X_2 &= Y_2 + Y_1 \\ Y_3 &= X_3 - \overline{X} & X_3 &= Y_3 + Y_1 \\ \vdots &= \quad \vdots & \vdots &= \quad \vdots \\ Y_n &= X_n - \overline{X} & X_n &= Y_n + Y_1 \end{aligned}$$

The Jacobian of the transformation can be shown to not depend on X_i or $\overline{X}$ and is equal to the constant n. It follows that

$$\begin{aligned} f_{Y_1, Y_2, \cdots, Y_n}(y_1, y_2, \cdots, y_n) &= f_X(x_1, x_2, \cdots, x_n)|J| \\ &= n f_X(x_1, y_1 + y_2, \cdots, y_1 + y_n) \\ &= \text{constants} \cdot \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2\right\} \qquad (3) \end{aligned}$$

Note that the sum in the exponent of the joint pdf can be simplified using Lemma 2. It follows that

$$\begin{aligned} \sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \\ &= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2\right] \\ &= \frac{1}{\sigma^2}\left[(x_1 - \bar{x})^2 + \sum_{i=2}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2\right] \qquad (4) \end{aligned}$$

Note that since $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$, it follows that

$$x_1 - \bar{x} = -\sum_{i=2}^{n}(x_i - \bar{x})$$

Therefore, equation (4) simplifies to

$$\begin{aligned}
\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 &= \frac{1}{\sigma^2}\left[(x_1-\bar{x})^2+\sum_{i=2}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu)^2\right] \\
&= \frac{1}{\sigma^2}\left[\left(\sum_{i=2}^{n}(x_i-\bar{x})\right)^2+\sum_{i=2}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu)^2\right] \\
&= \frac{1}{\sigma^2}\left[\left(\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2+n(y_1-\mu)^2\right]
\end{aligned}$$

Therefore, the pdf of $Y_1, Y_2, \cdots, Y_n$, equation (1), simplifies to

$$\begin{aligned}
f_{Y_1,Y_2,\cdots,Y_n}(y_1,y_2,\cdots,y_n) &= \text{constants}\cdot\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right\} \\
&= \text{constants}\cdot\exp\left\{-\frac{1}{2\sigma^2}\left[\left(\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2+n(y_1-\mu)^2\right]\right\} \\
&= \text{constants}\cdot\underbrace{\exp\left\{-\frac{1}{2\sigma^2}\left[\left(\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2\right]\right\}}_{h(y_2,y_3,\cdots,y_n)}\underbrace{\exp\left\{-\frac{n}{2\sigma^2}(y_1-\mu)^2\right\}}_{g(y_1)} \\
&= \text{constants}\cdot h(y_2,y_3,\cdots,y_n)\cdot g(y_1)
\end{aligned}$$

Because $f_{Y_1,Y_2,\cdots,Y_n}(y_1,y_2,\cdots,y_n)$ can be factored into a product of functions that depend only their respective set of statistics, it follows that $Y_1=\overline{X}$ is independent of $Y_i=X_i-\overline{X}$, $i=2,3,\cdots,n$.

Finally, since $X_1-\overline{X}=-\sum_{i=2}^{n}(X_i-\overline{X})$, it follows that $X_1-\overline{X}$ is a function of $X_i-\overline{X}$, $i=2,3,\cdots,n$. Therefore, $X_1-\overline{X}$ is independent of $Y_1=\overline{X}$. □

Theorem 2. *Suppose $X_1, X_2, \cdots, X_n$ is a random sample from a normal distribution with mean, μ, and variance, σ^2. It follows that the sample mean, $\overline{X}$, is independent of the sample variance, S^2.*

Proof. The definition of S^2 is given in Definition 1. Because S^2 is a function of $X_i-\overline{X}$, $i=1,2,\cdots,n$, it follows that S^2 is independent of $\overline{X}$. □

Theorem 3. *Suppose $X_1, X_2, \cdots, X_n$ is a random sample from a normal distribution with mean, μ, and variance, σ^2. It follows that the distribution of a mulitiple of the sample variance follows a χ^2 distribution with $n-1$ degrees of freedom. In particular,*

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

Proof. Equation (2) states

$$\sum_{i=1}^{n}(X_i-\mu)^2=(n-1)S^2+n(\overline{X}-\mu)^2.$$

It follows that

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} + n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2$$

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{(n-1)S^2}{\sigma^2} + \left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

$$U = W + V$$

Note that since $X_i \sim \mathrm{N}\big(\mu, \sigma^2\big)$, it follows that $Z_i = \dfrac{X_i-\mu}{\sigma} \sim \mathrm{N}(0,1)$. Similarly, since $\overline{X} \sim \mathrm{N}\Big(\mu, \frac{\sigma^2}{n}\Big)$, then $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$. By Lemma 3, it follows that $\left(\dfrac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1)$ and $V = \left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$. By Lemma 4, it follows that $U = \displaystyle\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$. Therefore, since W and V are independent, then the moment generating function of U is

$$M_U(t) = M_W(t)M_V(t)$$

$$(1-2t)^{-n/2} = M_W(t)(1-2t)^{-1/2}$$

$$\implies M_W(t) = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-(n-1)/2}$$

The moment generating function for W is recognized as coming from a χ^2 distribution with $n-1$ degrees of freedom. Hence,

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$\square$