Beginning of Sept.8, 2021

Remark. S is *not* Gaussian. To find its distribution, one way is to notice that both S and S^2 are positive, so

$$\mathbb{P}(S \leqslant t) = \mathbb{P}(S^2 \leqslant t^2) = \mathbb{P}((n-1)S^2/\sigma^2 \leqslant (n-1)t^2/\sigma^2) = \mathbb{P}(\chi^2_{n-1} \leqslant (n-1)t^2/\sigma^2) = \int_0^{(n-1)t^2/\sigma^2} f_{\chi^2_{n-1}}(x)\,\mathrm{d}x.$$

Differentiating the above expression (along with chain rule) would give us the pdf of S.

0.1 Student's *t*-distribution

Recall that

$$\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$$

has mean 0 and variance 1, given $X_1, \ldots, X_n$ are i.i.d. with finite mean and variance in $(0, \infty)$. Dividing both the numerator and the denominator by n gives

$$\frac{(X_1 + \ldots + X_n)/n - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}.$$

Now suppose that μ, σ are unknown, and we want to find them using $X_1, \ldots, X_n$.

It may be annoying to have two unknowns in one such equation, so sometimes we replace σ by the sample standard deviation, S, so that μ is the only free parameter, despite the fact that we don't know σ either:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}}.$$

If $X_1, \ldots, X_n$ are i.i.d. *Gaussians*, then the above quotient has the **Student's *t*-distribution**.

Proposition: (3.7) Student's t-distribution

Let X be a standard Gaussian random variable. Let Y be a chi-squared random variable with p degrees of freedom. Assume X and Y are independent. Then

$$\frac{X}{\sqrt{Y/p}}$$

has a **student's *t*-distribution** with

$$f_{X/\sqrt{Y/p}}(t) = \frac{\Gamma((p+1)/2)}{\sqrt{p\pi}\,\Gamma(p/2)}\left(1 + \frac{t^2}{p}\right)^{-(p+1)/2}, \qquad t \in \mathbb{R}.$$

Proof. First we define $Z := \sqrt{Y/p}$. It follows that for any $y > 0$,

$$\begin{aligned} f_Z(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\mathbb{P}(Z \leqslant y) = \frac{\mathrm{d}}{\mathrm{d}y}\mathbb{P}(Y \leqslant y^2 p) = \frac{\mathrm{d}}{\mathrm{d}y} \\ &= \frac{\mathrm{d}}{\mathrm{d}y}\int_0^{y^2 p} \frac{x^{p/2-1}e^{-x/2}}{2^{p/2}\Gamma(p/2)}\,\mathrm{d}x \\ &= 2yp \cdot p^{p/2-1} y^{p-2} e^{-y^2 p/2} \cdot \frac{1}{2^{p/2}\Gamma(p/2)} \\ &= 2y^{p-1} p^{p/2} e^{-y^2 p/2} \cdot \frac{1}{2^{p/2}\Gamma(p/2)}. \end{aligned}$$

Now we look at the CDF of X/Z:

$$\begin{aligned}\mathbb{P}(X/Z \leqslant t) = \mathbb{P}(X \leqslant tZ) &= \iint_{\substack{x \leqslant ty \\ y>0}} f_{X,Z}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ [\text{independence}] &= \iint_{\dots} f_X(x) f_Z(y)\,\mathrm{d}x\mathrm{d}y. \qquad (\Delta)\end{aligned}$$

Now we apply change of variable $\varphi\mathbb{R}^2 \to \mathbb{R}^2$ by $\varphi(a,b) = (ab, a)$ and $\varphi^{-1}(x,y) = (y, x/y)$. Then

$$|\mathcal{J}(a,b)| = \text{abs value of } \begin{vmatrix} b & 1 \\ a & 0 \end{vmatrix} = |a|.$$

Then (Δ) becomes

$$\begin{aligned}\iint_{\dots} f(x,y)\,\mathrm{d}x\mathrm{d}y &= \iint_{\varphi^{-1}(\dots)} f(\varphi(a,b))|\mathcal{J}(,)|\,\mathrm{d}a\mathrm{d}b \\ &= \iint_{\substack{b \leqslant t \\ a>0}} |a| f_X(ab) f_Z(b)\,\mathrm{d}a\mathrm{b} \\ &= \int_{b=-\infty}^{b=t} \int_{a=0}^{a=\infty} |a| f_X(ab) f_Z(b)\,\mathrm{d}a\,\mathrm{d}b. \qquad (\square)\end{aligned}$$

Recall that we will eventually $\mathrm{d}/\mathrm{d}t$ everything — this is exactly why we want $b = t$ as the upper limit of the outer integral: the derivative of this integral becomes

$$\begin{aligned}f_{X/Z}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(X/Z \leqslant t) \\ &= \int_{a=0}^{a=\infty} |a| f_X(at) f_Z(b)\,\mathrm{d}a \\ &= \frac{p^{p/2}}{\sqrt{2\pi}} \int_0^\infty a e^{-a^2t^2/2} a^{p-1} e^{-a^2t^2/2}\,\mathrm{d}x\end{aligned}$$

$\square$