1. Given independent random variables $X_1, \ldots, X_n$, define

$$W = \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n X_i^2}} = \frac{n^{1/2} \bar{X}}{\sqrt{S^2/n}}, \quad T = \frac{n^{1/2} \bar{X}}{S},$$

where $S^2 = (n-1) \sum_{i=1}^n (X_i - \bar{X})^2$. Show that the following identity holds

$$T = \left( \frac{n-1}{n} \right)^{1/2} \frac{W}{\sqrt{1 - W^2/n}}.$$

And $W$ and $T$ have a one-to-one correspondence.

Solution: The identity can be proved by a direct computation,

$$T = \frac{\sum_{i=1}^n X_i}{n^{1/2} S} = \left( \frac{n-1}{n} \right)^{1/2} \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n X_i^2}} \frac{1}{\sqrt{1 - n X^2/\sum_{i=1}^n X_i^2}} = \left( \frac{n-1}{n} \right)^{1/2} \frac{W}{\sqrt{1 - W^2/n}}.$$

By Cauchy–Schwarz inequality, $\sum_{i=1}^n X_i^2 \geq n\bar{X}^2$, so $|W| < \sqrt{n}$, and the one-to-one correspondence follows from the fact that $f(x) := x/\sqrt{1 - x^2/n}$ is strictly increasing over $x \in (-\sqrt{n}, \sqrt{n})$.

2. Let $U_1/\sigma_1^2 \sim \chi_{d_1}^2$, and $U_2/\sigma_2^2 \sim \chi_{d_2}^2$, and they are independent. Suppose $\sigma_2^2/\sigma_1^2 = a$. Show that $U_2/U_1$ and $aU_1 + U_2$ are independent. In particular, if $\sigma_1 = \sigma_2$, $U_2/U_1$ and $U_1 + U_2$ are independent.

Solution: The density of $U_1$ is

$$f_{U_1}(u_1) = \frac{1}{2^{d_1/2} \Gamma(d_1/2)} u_1^{d_1/2 - 1} \sigma_1^{-d_1} e^{-u_1/(2\sigma_1^2)},$$

and the density of $U_2$ can be similarly calculated. The joint density of $(U_1, U_2)$ is then proportional to

$$f_{(U_1, U_2)}(u_1, u_2) \propto \exp(-(au_1 + u_2)/(2\sigma_2^2)),$$

so $aU_1 + U_2$ is sufficient and (boundedly) complete. Then, notice that

$$\frac{U_2}{U_1} = \frac{\sigma_2^2}{\sigma_1^2} \frac{U_2/\sigma_2^2}{U_1/\sigma_1^2} \sim aF_{d_2, d_1},$$

which is independent of $\sigma_2$. Applying Basu’s theorem gives us the desired independence.

3. Suppose that random vector $(X, Y)$ has probability density function

$$\frac{1}{\pi} e^{-\frac{x^2 + y^2}{2}} \mathbb{1}(xy > 0), \quad x, y \in \mathbb{R}.$$

Does $(X, Y)$ possess a multivariate normal distribution? Find the marginal distributions.

Solution: No. For example, $\mathbb{P}(X < 0, Y > 0) = 0$, but it cannot be zero if $(X, Y)$ possess a bivariate normal distribution. Marginally, when $x > 0$,

$$f_X(x) = \int_0^\infty \frac{1}{\pi} e^{-\frac{x^2 + y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

when $x < 0$,

$$f_X(x) = \int_{-\infty}^0 \frac{1}{\pi} e^{-\frac{x^2 + y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

and $f_X(0) = 0$. The marginal distribution of $Y$ is the same.
4. Suppose that \( X_m \sim \text{Binomial}(m, p_1) \), \( Y_n \sim \text{Binomial}(n, p_2) \) and they are independent. To test \( H_0 : p_1 = p_2 = p \) for some predetermined \( p \in (0, 1) \), consider the test statistic

\[
C_{m,n}^2 = \frac{(X_m - mp)^2}{mp(1-p)} + \frac{(Y_n - np)^2}{np(1-p)}.
\]

(a) Find the limit distribution of \( C_{m,n}^2 \) as \( m, n \to \infty \);

**Solution:** We can write \( X_m = \sum_{i=1}^{m} \tilde{X}_i \) and \( Y_n = \sum_{i=1}^{n} \tilde{Y}_i \), where \( \tilde{X}_i \sim \text{Bernoulli}(p_1) \) and \( \tilde{Y}_i \sim \text{Bernoulli}(p_2) \). Under null hypothesis \( p_1 = p_2 = p \), we have

\[
\frac{X_m - mp}{\sqrt{mp(1-p)}} \xrightarrow{d} N(0, 1), \quad \frac{Y_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1)
\]
from central limit theorem. This leads us to the conclusion

\[
C_{m,n}^2 \xrightarrow{d} \chi^2_2.
\]

(b) How would you modify the test statistic if \( p \) were unknown? What is the limit distribution after modification?

**Solution 1:** Consider the two-sample proportions test, and the test statistic is

\[
z = \frac{X_m/m - Y_n/n}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right) \cdot \frac{X_m + Y_n}{m + n} \cdot \left(1 - \frac{X_m + Y_n}{m + n}\right)}}.
\]

The limit distribution is \( \mathcal{N}(0, 1) \).

**Solution 2:** In the original statistic \( C_{m,n}^2 \), we replace \( p \) with its MLE

\[
\hat{p} = \frac{X_m + Y_n}{m + n}.
\]

By doing this, we add one more restriction so that one degree of freedom is sacrificed, and the limit distribution becomes \( \chi^2_1 \).

In fact, the above two tests are equivalent.