On the Distribution of γ When ρ is not Zero

On the Distribution of r When ρ is not Zero

Masahiro ŌBA

The Study of Mathematics, Iwamizawa Branch, Hokkaido University of Education

大場将寛 ： ρ がゼロでないときの r の分布について

Although the exact distribution of r for samples from a correlated parent population was found by Fisher, using a geometrical method,[4] this method is very difficult. Sawkins has given us its analitical treatment. But this was a special case, where the means of variates were zero. The following states the most general case, on the supposition that two variates are normal and linear, giving an orthogonal linear transformation.

The linear and most general probability density function is given by the following

$$p(x, y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\xi_1)^2}{\sigma_1^2}-\frac{2\rho(x-\xi_1)(y-\xi_2)}{\sigma_1\sigma_2}+\frac{(y-\xi_2)^2}{\sigma_2^2}\right\}}$$

Writing $t=\frac{x-\xi_1}{\sigma_1}$, $u=\frac{y-\xi_2}{\sigma_2}$, the joint frequency distribution of t and u is

$$p(t, u)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\{t^2-2\rho tu+u^2\}}$$

$$=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\{(u-\rho t)^2+(1-\rho^2)t^2\}}$$

Hence, if $v=\frac{u-\rho t}{\sqrt{1-\rho^2}}$ the joint distribution of t and v is given by

$$p(t, v)=\frac{1}{2\pi}e^{-\frac{v^2}{2}-\frac{t^2}{2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}.$$

t and v are therefore independent normal standard variates, and $\sum_{i=1}^{n}v_i^2$ is distributed as χ^2 with n degrees of freedom.

By the following transformation from the v_i to new variates z_i,

$$v_1=\frac{1}{\sqrt{n}}z_1+\sqrt{\frac{1}{1\cdot2}}z_2+\sqrt{\frac{1}{2\cdot3}}z_3+\sqrt{\frac{1}{3\cdot4}}z_4+\cdots\cdots+\sqrt{\frac{1}{(n-1)n}}z_n$$

$$v_2=\frac{1}{\sqrt{n}}z_1-\sqrt{\frac{1}{1\cdot2}}z_2+\sqrt{\frac{1}{2\cdot3}}z_3+\sqrt{\frac{1}{3\cdot4}}z_4+\cdots\cdots+\sqrt{\frac{1}{(n-1)n}}z_n$$

$$v_3 = \frac{1}{\sqrt{n}}z_1 \qquad -2\sqrt{\frac{1}{2\cdot 3}}z_3 + \sqrt{\frac{1}{3\cdot 4}}z_4 + \cdots\cdots + \sqrt{\frac{1}{(n-1)n}}z_n$$
$$v_4 = \frac{1}{\sqrt{n}}z_1 \qquad -3\sqrt{\frac{1}{3\cdot 4}}z_4 + \cdots\cdots + \sqrt{\frac{1}{(n-1)n}}z_n$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$v_n = \frac{1}{\sqrt{n}}z_1 \qquad -(n-1)\sqrt{\frac{1}{(n-1)n}}z_n,$$

we have $\sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n} z_i^2$. Of course, z_i are normal standard variates, and $\sum_{i=1}^{n} z_i^2$ is distributed as χ^2 with n degrees of freedom, and this transformation is an orthogonal linear transformation.

Choosing $z_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} v_i$, we have

$$z_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{u_i-\rho t_i}{\sqrt{1-\rho^2}} = \frac{1}{\sqrt{n}}\,\frac{\sum_{i=1}^{n}u_i - \rho\sum_{i=1}^{n}t_i}{\sqrt{1-\rho^2}}$$
$$= \frac{1}{\sqrt{n}}\,\frac{n\bar{u}-\rho n\bar{t}}{\sqrt{1-\rho^2}} = \frac{\sqrt{n}\,(\bar{u}-\rho\bar{t})}{\sqrt{1-\rho^2}},$$

and

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} v_i^2 = \sum_{i=1}^{n}\frac{u_i^2-2\rho u_i t_i+\rho^2 t_i^2}{1-\rho^2} = \frac{1}{1-\rho^2}\{\sum_{i=1}^{n}u_i^2 - 2\rho\sum_{i=1}^{n}u_i t_i + \rho^2\sum_{i=1}^{n}t_i^2\}$$
$$= \frac{1}{1-\rho^2}\{\sum_{i=1}^{n}(u_i-\bar{u})^2 - 2\rho\sum_{i=1}^{n}(u_2-\bar{u})(t_i-\bar{t}) + \rho^2\sum_{i=1}^{n}(t_i-\bar{t})^2$$
$$+ n(\bar{u}^2 - 2\rho\bar{u}\bar{t} + \rho^2\bar{t}^2)\}$$
$$= \frac{1}{1-\rho^2}\{\sum_{i=1}^{n}(u_i-\bar{u})^2 - 2\rho\sum_{i=1}^{n}(u_i-\bar{u})(t_i-\bar{t}) + \rho^2\sum_{i=1}^{n}(t_i-\bar{t})^2 + n(\bar{u}-\rho\bar{t})^2\}$$
$$= \frac{1}{1-\rho^2}\{S_2^2 - 2\rho r S_2 S_1 + \rho^2 S_1^2\} + z_1^2,$$

where we put $S_2^2 = \sum_{i=1}^{n}(u_i-\bar{u})^2$, $S_1^2 = \sum_{i=1}^{n}(t_i-\bar{t})^2$, $r = \frac{\sum_{i=1}^{n}(u_i-\bar{u})(t_i-\bar{t})}{S_2 S_1}$.

Therefore we have

$$\sum_{i=2}^{n} z_i^2 = \frac{1}{1-\rho^2}\{S_2^2 - 2\rho r S_2 S_1 + \rho^2 S_1^2\},$$

and this is distributed as χ^2 with $n-1$ degrees of freedom.

Choosing $z_2 = \frac{\sum_{i=1}^{n}(t_i-\bar{t})v_i}{S_1}$, we have

$$z_2 = \frac{\sum_{i=1}^{n}(t_i-\bar{t})(u_i-\rho t_i)}{S_1\sqrt{1-\rho^2}} = \frac{\sum_{i=1}^{n}(t_i-\bar{t})(u_i-\rho t_i) - (n\bar{t}-n\bar{t})(\bar{u}-\rho\bar{t})}{S_1\sqrt{1-\rho^2}}$$

$$=\frac{\sum_{i=1}^{n}(t_i-\bar{t})(u_i-\rho t_i)-\sum_{i=1}^{n}(t_i-\bar{t})(\bar{u}-\rho\bar{t})}{S_1\sqrt{1-\rho^2}}=\frac{\sum_{i=1}^{n}(t_i-\bar{t})\{(u_i-\rho t_i)-(\bar{u}-\rho\bar{t})\}}{S_1\sqrt{1-\rho^2}}$$

$$=\frac{\sum_{i=1}^{n}(t_i-\bar{t})\{(u_i-\bar{u})-\rho(t_i-\bar{t})\}}{S_1\sqrt{1-\rho^2}}=\frac{\sum_{i=1}^{n}(t_i-\bar{t})(u_i-\bar{u})-\rho\sum_{i=1}^{n}(t_i-\bar{t})^2}{S_1\sqrt{1-\rho^2}}$$

$$=\frac{1}{\sqrt{1-\rho^2}}\left\{\frac{\sum_{i=1}^{n}(t_i-\bar{t})(u_i-\bar{u})}{S_1}-\rho\frac{\sum_{i=1}^{n}(t_i-\bar{t})^2}{S_1}\right\}=\frac{1}{\sqrt{1-\rho^2}}\{rS_2-\rho S_1\}.$$

Therefore we have

$$z_2^2=\frac{r^2S_2^2-2r\rho S_2S_1+\rho^2S_1^2}{1-\rho^2}$$

and

$$\sum_{i=3}^{n}z_i^2=\sum_{i=2}^{n}z_i^2-z_2^2=\frac{1}{1-\rho^2}\{S_2^2-2\rho rS_2S_1+\rho^2S_1^2\}-\frac{r^2S_2^2-2r\rho S_2S_1+\rho^2S_1^2}{1-\rho^2}=\frac{S_2^2(1-r^2)}{1-\rho^2}.$$

This is distributed as χ^2 with $n-2$ degrees of freedom.

Moreover, as $S_1^2=\sum_{i=1}^{n}(t_i-\bar{t})^2$ is independently distributed as χ^2 with $n-1$ degrees of freedom, if we name

$$a=z_2=\frac{rS_2-\rho S_1}{\sqrt{1-\rho^2}},$$

$$b=\frac{1}{2}\sum_{i=3}^{n}z_i^2=\frac{S_2^2(1-r^2)}{2(1-\rho^2)},$$

$$c=\frac{1}{2}\sum_{i=1}^{n}(t_i-\bar{t})^2=\frac{1}{2}S_1^2,$$

a, b, and c are respectively independent, a is $N(0,\ 1)$ distribution, $2b$ is distributed as χ^2 with $n-2$ degrees of freedom, and $2c$ is distributed as χ^2 with $n-1$ degrees of freedom.

Using the above relations, we have

$$f(a,\ b,\ c)=f(a)f(b)f(c)=f(a)\cdot 2f(2b)\cdot 2f(2c).$$

As

$$f(b)=2\cdot\frac{1}{2^{\frac{n-2}{2}}\Gamma\left(\frac{n-2}{2}\right)}(2b)^{\frac{n-2}{2}-1}e^{-\frac{2b}{2}}=\frac{1}{\Gamma\left(\frac{n-2}{2}\right)}b^{\frac{n-4}{2}}e^{-b},$$

$$f(c)=2\cdot\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}(2c)^{\frac{n-1}{2}-1}e^{-\frac{2c}{2}}=\frac{1}{\Gamma\left(\frac{n-1}{2}\right)}c^{\frac{n-3}{2}}e^{-c},$$

we have

$$f(a,\ b,\ c)=\frac{1}{\sqrt{2\pi}}e^{-\frac{a^2}{2}}\cdot\frac{1}{\Gamma\left(\frac{n-2}{2}\right)}b^{\frac{n-4}{2}}e^{-b}\cdot\frac{1}{\Gamma\left(\frac{n-1}{2}\right)}c^{\frac{n-3}{2}}e^{-c}$$

$$=e^{-\left(\frac{a^2}{2}+b+c\right)}\left\{\frac{S_2^2(1-r^2)}{2(1-\rho^2)}\right\}^{\frac{n-4}{2}}\left\{\frac{S_1^2}{2}\right\}^{\frac{n-3}{2}}\frac{1}{\sqrt{2\pi}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}.$$

On the other hand, the relations

$$\frac{a^2}{2}+b+c=\frac{(rS_2-\rho S_1)^2}{2(1-\rho^2)}+\frac{S_2^2(1-r^2)}{2(1-\rho^2)}+\frac{S_1^2(1-\rho^2)}{2(1-\rho^2)}=\frac{S_1^2+S_2^2-2\rho rS_1S_2}{2(1-\rho^2)},$$

$$\left\{\frac{S_2^2(1-r^2)}{2(1-\rho^2)}\right\}^{\frac{n-4}{2}}\left\{\frac{S_1^2}{2}\right\}^{\frac{n-3}{2}}=\frac{S_1^{\ n-3}S_2^{\ n-4}(1-r^2)^{\frac{n-4}{2}}}{2^{\frac{2n-7}{2}}(1-\rho^2)^{\frac{n-4}{2}}}$$

are constant, so we have

$$f(a,\ b,\ c)\ =e^{-\frac{S_1^2+S_2^2-2\rho rS_1S_2}{2(1-\rho^2)}}\frac{(1-r^2)^{\frac{n-4}{2}}S_1^{\ n-3}S_2^{\ n-4}}{(1-\rho^2)^{\frac{n-4}{2}}2^{\frac{2n-7}{2}}\sqrt{2\pi}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}$$

Next, we must find a probability density function $f(r,\ S_1,\ S_2)$.
That is

$$f(r,\ S_1,\ S_2)=f(a,\ b,\ c)\cdot\left|J\left(\frac{a,\ b,\ c}{r,\ S_1,\ S_2}\right)\right|=f(a,\ b,\ c)\frac{S_1S_2^2}{(1-\rho^2)^{\frac{3}{2}}}=e^{-\frac{S_1^2+S_2^2-2\rho rS_1S_2}{2(1-\rho^2)}}$$

$$\times\frac{(1-r^2)^{\frac{n-4}{2}}S_1^{\ n-2}S_2^{\ n-2}}{(1-\rho^2)^{\frac{n-1}{2}}\sqrt{\pi}\ 2^{n-3}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)},$$

because of

$$J\left(\frac{a,\ b,\ c}{r,\ S_1,\ S_2}\right)=\begin{vmatrix}\frac{S_2}{\sqrt{1-\rho^2}} & \frac{-\rho}{\sqrt{1-\rho^2}} & \frac{r}{\sqrt{1-\rho^2}}\\ \frac{-rS_2^2}{1-\rho^2} & 0 & \frac{S_2(1-r^2)}{1-\rho^2}\\ 0 & S_1 & 0\end{vmatrix}=\frac{1}{(1-\rho^2)^{\frac{3}{2}}}\begin{vmatrix}S_2 & -\rho & r\\ -rS_2^2 & 0 & S_2(1-r^2)\\ 0 & S_1 & 0\end{vmatrix}$$

$$=\frac{r(-rS_2^2)S_1-S_2S_2(1-r^2)S_1}{(1-\rho^2)^{\frac{3}{2}}}=\frac{-S_1S_2^2}{(1-\rho^2)^{\frac{3}{2}}}.$$

The frequency function for r is given by integrating with respect to S_1 and S_2 from 0 to ∞. This integration is not so easy. Fisher found an ingenious transformation from S_1 and S_2 to new variables α and β such as

$$S_1=\sqrt{\alpha}\,e^{\frac{\beta}{2}},\qquad S_2=\sqrt{\alpha}\,e^{-\frac{\beta}{2}}.$$

By the following relation

$$S_1^2+S_2^2-2\rho r S_1S_2=(\sqrt{\alpha}\,e^{\frac{\beta}{2}})^2+(\sqrt{\alpha}\,e^{-\frac{\beta}{2}})^2-2\rho r(\sqrt{\alpha}\,e^{\frac{\beta}{2}})(\sqrt{\alpha}\,e^{-\frac{\beta}{2}})$$
$$=2\alpha\left(\frac{e^{\beta}+e^{-\beta}}{2}\right)-2\rho r\alpha=2\alpha(\cos h\beta-\rho r),$$

we have

$$f(r,\ S_1,\ S_2)=e^{-\frac{\alpha(\cos h\beta-\rho r)}{1-\rho^2}}\frac{(1-r^2)^{\frac{n-4}{2}}\alpha^{n-2}}{(1-\rho^2)^{\frac{n-1}{2}}\sqrt{\pi}\,2^{n-3}\,\Gamma\left(\frac{n-1}{1}\right)\Gamma\left(\frac{n-2}{2}\right)}.$$

Next, moreover we must find a function $f(r,\ \alpha,\ \beta)$.
we have

$$f(r,\ \alpha,\ \beta)=f(r,\ S_1,\ S_2)\left|J\left(\frac{S_1,\ S_2}{\alpha,\ \beta}\right)\right|=f(r,\ S_1,\ S_2)\frac{1}{2}$$
$$=\frac{1}{2}e^{-\frac{\alpha(\cos h\beta-\rho r)}{1-\rho^2}}\frac{(1-r^2)^{\frac{n-4}{2}}\alpha^{n-2}}{(1-\rho^2)^{\frac{n-1}{2}}\sqrt{\pi}\,2^{n-3}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)},$$

because of

$$J\left(\frac{S_1,\ S_2}{\alpha,\ \beta}\right)=\begin{vmatrix}\frac{1}{2}\alpha^{-\frac{1}{2}}e^{\frac{\beta}{2}} & \frac{1}{2}\alpha^{\frac{1}{2}}e^{\frac{\beta}{2}}\\ \frac{1}{2}\alpha^{-\frac{1}{2}}e^{-\frac{\beta}{2}} & -\frac{1}{2}\alpha^{\frac{1}{2}}e^{-\frac{\beta}{2}}\end{vmatrix}=-\frac{1}{2}.$$

Seeing that the limits of α are from 0 to ∞ and those of β from $-\infty$ to ∞, we have

$$f(r,\ \beta)=\int_0^{\infty}\frac{1}{2}e^{-\frac{\alpha(\cos h\beta-\rho r)}{1-\rho^2}}\frac{(1-r^2)^{\frac{n-4}{2}}\alpha^{n-2}}{(1-\rho^2)^{\frac{n-1}{2}}\sqrt{\pi}\,2^{n-3}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}d\alpha.$$

When we put $\dfrac{\alpha(\cos h\beta-\rho r)}{1-\rho^2}=\gamma$, we have

$$\alpha^{n-2}=\frac{(1-\rho^2)^{n-2}}{(\cos h\beta-\rho r)^{n-2}}\gamma^{n-2}\ ,\quad d\alpha=\frac{1-\rho^2}{\cos h\beta-\rho r}d\gamma,$$

and then we have

$$f(r,\ \beta)=\frac{1(1-r^2)^{\frac{n-4}{2}}}{2(1-\rho^2)^{-\frac{n-1}{2}}\sqrt{\pi}\,2^{n-3}\ \Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)(\cos h\beta-\rho r)^{n-1}}$$

$$\times\int_0^{\infty}e^{-\gamma}\ \gamma^{n-2}\ d\gamma$$

$$=\frac{(1-r^2)^{\frac{n-4}{2}}(1-\rho^2)^{\frac{n-1}{2}}\ \Gamma(n-1)}{2\sqrt{\pi}\,2^{n-3}\ \Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)(\cos h\beta-\rho r)^{n-1}}.$$

On the other hand, $2^{n-3}\ \Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)=\sqrt{\pi}\ \Gamma(n-2)$ is constant owing to

$$\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)=\frac{n-3}{2}\ \frac{n-4}{2}\ \frac{n-5}{2}\ \frac{n-6}{2}\ \frac{n-7}{2}\ \frac{n-8}{2}\cdots\cdots\cdots\cdots\frac{3}{2}\ \frac{2}{2}\ \frac{1}{2}\Gamma\left(\frac{1}{2}\right)$$

$$=\frac{1}{2^{n-3}}(n-3)(n-4)(n-5)(n-6)\cdots\cdots\cdots3\cdot2\cdot1\cdot\sqrt{\pi}$$

$$=\frac{1}{2^{n-3}}\Gamma(n-2)\sqrt{\pi},$$

therefore we have

$$f(r,\ \beta)=\frac{(1-r^2)^{\frac{n-4}{2}}(1-\rho^2)^{\frac{n-1}{2}}\ \Gamma(n-1)}{2\pi\Gamma(n-2)(\cos h\beta-\rho r)^{n-1}}=\frac{n-2}{2\pi}(1-r^2)^{\frac{n-4}{2}}$$

$$\times(1-\rho^2)^{\frac{n-1}{2}}\frac{1}{(\cos h\beta-\rho r)^{n-1}}.$$

Finally we get

$$f(r)=\int_{-\infty}^{\infty}f(r,\ \beta)d\beta=2\int_0^{\infty}f(r,\ \beta)=\frac{n-2}{\pi}(1-r^2)^{\frac{n-4}{2}}$$

$$\times(1-\rho^2)^{\frac{n-1}{2}}\int_0^{\infty}\frac{d\beta}{(\cos h\beta-\rho r)^{n-1}}.$$

References

〔1〕 Anderson, T. W. (1958) Introduction to Multivariate Statistical Analysis. Wiley, New York. p. 66–74.
〔2〕 Fisher, R. A. (1915) Biometrika. Cambridge–London. Vol. 10, p. 507.
〔3〕 Fisher, R. A. (1921) Metron. Ferare–Rovigo–Roma. Vol. 1, pt. 4, p. 14.
〔4〕 Kenney, J. F. (1951) Mathmatics of Statistics. Van Nostrand Company. pt. 2, p. 217–219.
〔5〕 Sawkins, D. T. (1944) J. and Proc. R. Soc. N. S. W. Vol. 77, p. 85–95.