Skellam distribution

The **Skellam distribution** is the discrete probability distribution of the difference $n_1 - n_2$ of two statistically independent random variables N_1 and N_2 each having Poisson distributions with different expected values μ_1 and μ_2 . It is useful in describing the statistics of the difference of two images with simple photon noise, as well as describing the point spread distribution in sports where all scored points are equal, such as baseball, hockey and soccer.

The distribution is also applicable to a special case of the difference of dependent Poisson random variables, but just the obvious case where the two variables have a common additive random contribution which is cancelled by the differencing: see Karlis & Ntzoufras (2003) for details and an application.

The probability mass function for the Skellam distribution for a count difference $k = n_1 - n_2$ from two Poisson-distributed variables with means μ_1 and μ_2 is given by:

$$f(k;\mu_1,\mu_2) = e^{-(\mu_1+\mu_2)} \left(\frac{\mu_1}{\mu_2}\right)^{k/2} I_k(2\sqrt{\mu_1\mu_2})$$

where *Ik*(*z*) is the modified Bessel function of the first kind. Note that since *k* is an integer we have that *Ik*(*z*)=*I*|*k*|(*z*).

1 Derivation

Note that the probability mass function of a Poisson distribution for a count *n* with mean μ is given by

$$f(n;\mu) = \frac{\mu^n}{n!}e^{-\mu}.$$

for $n \geq 0$ (and zero otherwise). The Skellam probability mass function for the difference of two counts $k = n_1 - n_2$ is the cross-correlation of two Poisson distributions: (Skellam, 1946)

$$f(k;\mu_1,\mu_2) = \sum_{n=-\infty}^{\infty} f(k+n;\mu_1)f(n;\mu_2)$$

$$= e^{-(\mu_1+\mu_2)} \sum_{n=max(0,-k)}^{\infty} \frac{\mu_1^{k+n}\mu_2^n}{n!(k+n)!}$$

Since the Poisson distribution is zero for negative values of the count $(f(n<0;\mu)=0)$, the second sum is only taken for those terms where $n >= 0$ and $n+k >= 0$. It can be shown that the above sum implies that

$$\frac{f(k;\mu_1,\mu_2)}{f(-k;\mu_1,\mu_2)} = \left(\frac{\mu_1}{\mu_2}\right)^k$$

so that:

$$f(k;\mu_1,\mu_2) = e^{-(\mu_1+\mu_2)} \left(\frac{\mu_1}{\mu_2}\right)^{k/2} I_{|k|}(2\sqrt{\mu_1\mu_2})$$

where *I* $_k$(z) is the modified Bessel function of the first kind. The special case for $\mu_1 = \mu_2 (= \mu)$ is given by Irwin (1937):

$$f\left(k;\mu,\mu\right) = e^{-2\mu}I_{|k|}(2\mu).$$

Note also that, using the limiting values of the modified Bessel function for small arguments, we can recover the Poisson distribution as a special case of the Skellam distribution for $\mu_2 = 0$.

2 Properties

As it is a discrete probability function, the Skellam probability mass function is normalized:

$$\sum_{k=-\infty}^{\infty} f(k;\mu_1,\mu_2) = 1.$$

We know that the probability generating function (pgf) for a Poisson distribution is:

$$G\left(t;\mu\right) = e^{\mu(t-1)}.$$

It follows that the pgf, $G(t;\mu_1,\mu_2)$, for a Skellam probability function will be:

$$G(t;\mu_1,\mu_2) = \sum_{k=0}^{\infty} f(k;\mu_1,\mu_2)t^k$$

$$= G(t;\mu_1)\,G(1/t;\mu_2)$$
$$= e^{-(\mu_1+\mu_2)+\mu_1 t+\mu_2/t}.$$

Notice that the form of the probability generating function implies that the distribution of the sums or the differences of any number of independent Skellam-distributed variables are again Skellam-distributed. It is sometimes claimed that any linear combination of two Skellam-distributed variables are again Skellam-distributed, but this is clearly not true since any multiplier other than +/-1 would change the support of the distribution and alter the pattern of moments in a way that no Skellam distribution can satisfy.

The moment-generating function is given by:

$$M(t;\mu_1,\mu_2) = G(e^t;\mu_1,\mu_2)$$
$$= \sum_{k=0}^{\infty} \frac{t^k}{k!}\, m_k$$

which yields the raw moments *mk* . Define:

$$\Delta \stackrel{\text{def}}{=} \mu_1 - \mu_2$$
$$\mu \stackrel{\text{def}}{=} (\mu_1+\mu_2)/2.$$

Then the raw moments *mk* are

$$m_1 = \Delta$$
$$m_2 = 2\mu + \Delta^2$$
$$m_3 = \Delta(1 + 6\mu + \Delta^2)$$

The central moments *Mk* are

$$M_2 = 2\mu\ ,$$
$$M_3 = \Delta\ ,$$
$$M_4 = 2\mu + 12\mu^2\ .$$

The mean, variance, skewness, and kurtosis excess are respectively:

$$E(n) = \Delta$$
$$\sigma^2 = 2\mu$$
$$\gamma_1 = \Delta/(2\mu)^{3/2}$$
$$\gamma_2 = 1/2\mu\ .$$

The cumulant-generating function is given by:

$$K(t;\mu_1,\mu_2) \stackrel{\text{def}}{=} \ln(M(t;\mu_1,\mu_2)) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\,\kappa_k$$

which yields the cumulants:

$$\kappa_{2k} = 2\mu$$
$$\kappa_{2k+1} = \Delta\ .$$

For the special case when μ1 = μ2, an asymptotic expansion of the modified Bessel function of the first kind yields for large μ:

$$f(k;\mu,\mu) \sim \frac{1}{\sqrt{4\pi\mu}}\left[1 + \sum_{n=1}^{\infty}(-1)^n \frac{\{4k^2-1^2\}\{4k^2-3^2\}\cdots\{4k^2-(}{n!\,2^{3n}\,(2\mu)^n}\right.$$

(Abramowitz & Stegun 1972, p. 377). Also, for this special case, when *k* is also large, and of order of the square root of 2μ, the distribution tends to a normal distribution:

$$f(k;\mu,\mu) \sim \frac{e^{-k^2/4\mu}}{\sqrt{4\pi\mu}}.$$

These special results can easily be extended to the more general case of different means.

Recurrence relation

$$\left\{-\mu_1 P(k) + \mu_2 P(k+2) + (k+1)P(k+1) = 0, P(0) = e^{-\mu_1-\mu_2}\,{}_0\tilde{F}\right.$$

2.1 Bounds on weight above zero

If $X \sim Skellam(\mu_1,\mu_2)$, with $\mu_1 < \mu_2$, then

$$\frac{\exp(-(\sqrt{\mu_1}-\sqrt{\mu_2})^2)}{(\mu_1+\mu_2)^2} - \frac{e^{-(\mu_1+\mu_2)}}{2\sqrt{\mu_1\mu_2}} - \frac{e^{-(\mu_1+\mu_2)}}{4\mu_1\mu_2} \le P(X \ge 0) \le \exp(-$$

Details can be found in Poisson distribution#Poisson_Races

3 References

- Abramowitz, Milton; Stegun, Irene A., eds. (June 1965). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Unabridged and unaltered republ. [der Ausg.] 1964, 5. Dover printing ed.). Dover Publications. pp. 374–378. ISBN 0486612724. Retrieved 27 September 2012.

- Irwin, J. O. (1937) “The frequency distribution of the difference between two independent variates following the same Poisson distribution.” *Journal of the Royal Statistical Society: Series A*, 100 (3), 415–416.

- Karlis, D. and Ntzoufras, I. (2003) "Analysis of sports data using bivariate Poisson models". *Journal of the Royal Statistical Society, Series D*, 52 (3), 381–393. doi:10.1111/1467-9884.00366
- Karlis D. and Ntzoufras I. (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine*, 25, 1885–1905.
- Skellam, J. G. (1946) "The frequency distribution of the difference between two Poisson variates belonging to different populations". *Journal of the Royal Statistical Society, Series A*, 109 (3), 296.

4 Text and image sources, contributors, and licenses

4.1 Text

- **Skellam distribution** *Source:* http://en.wikipedia.org/wiki/Skellam%20distribution?oldid=646576484 *Contributors:* Michael Hardy, Tomi, SebastianHelm, Charles Matthews, Giftlite, Johnprovis, Jason Quinn, MarkSweep, John Vandenberg, PAR, Bkkbrad, Btyner, Rjwilmsi, Krishnavedala, Halcyonhazard, Mebden, Shorespirit, Skittleys, Thijs!bot, Magioladitis, Ankitdoshi1, JLewis, Melcombe, Borneq, Qwfp, Addbot, Alexradi, Luckas-bot, Xqbot, FrescoBot, Citing, Illia Connell, Gmk7, BeyondNormality, Monkbot and Anonymous: 15

4.2 Images

- **File:Skellam_distribution.svg** *Source:* http://upload.wikimedia.org/wikipedia/commons/a/a9/Skellam_distribution.svg *License:* CC0 *Contributors:* Own work *Original artist:* Krishnavedala

4.3 Content license