MATH 162: SUMS OF POISSON RANDOM VARIABLES

STEVEN J. MILLER

Abstract. We show that, appropriately scaled, the mean of n independent Poisson variables converges to the standard normal distribution $N(0, 1)$.

1. Review

Theorem 1.1. *Let X be a Poisson random variable with parameter λ. Its moment generating function satisfies*

$$M_X(t) \;=\; e^{\lambda(e^t-1)}. \tag{1.1}$$

Note the mean is $\mu_X = \lambda$ and the variance is $\sigma_X^2 = \lambda$.

Theorem 1.2. *Let X be a normal random variable with mean μ and variance σ^2. Its moment generating function satisfies*

$$M_X(t) \;=\; e^{\mu t+\frac{\sigma^2 t^2}{2}}. \tag{1.2}$$

2. Sum of Poisson Random Variables

Let X_i be Poisson random variables with parameter λ. Let

$$Y_n \;=\; X_1 + \cdots + X_n. \tag{2.1}$$

We expect Y_n to be of size $n\mu = n\lambda$. This follows from the linearity of expected value:

$$\mathbb{E}[Y_n] \;=\; \sum_i 1 \cdot \mathbb{E}[X_i] \;=\; \sum_i \lambda \;=\; n\lambda. \tag{2.2}$$

Let σ denote the variance of X (the Poisson distribution with parameter λ). The variance of Y_n is computed similarly; since the X_i are independent we have

$$\mathrm{Var}(Y_n) \;=\; \sigma_{Y_n}^2 \;=\; \sum_i 1^2 \cdot \mathrm{Var}(X_i) \;=\; n\sigma^2. \tag{2.3}$$

Note this is a little different than class because we have *not* divided Y_n by n; if $W_n = \frac{1}{n}Y_n$, then $\mathrm{Var}(W_n) = \frac{1}{n^2}\mathrm{Var}(Y_n) = \frac{\sigma^2}{n}$.

The natural quantity to study is

$$Z_n \;=\; \frac{Y_n - n\lambda}{\sigma_{Y_n}} \;=\; \frac{(X_1 + \cdots + X_n) - n\lambda}{\sigma\sqrt{n}}. \tag{2.4}$$

The reason this is the natural quantity is that the sum of the X_i is expected to be around $n\lambda$; if we subtract the predicted value, what is left is the fluctuations about the mean. We then need to figure out what are the correct units. As the variance of the sum of the X_is is $n\sigma^2$, its standard deviation is $\sigma\sqrt{n}$; thus, it is natural to measure the difference from the predicted mean in units of the expected standard deviation.

Note this is a little different than class (I accidentally wrote $\sigma/\sqrt{n}$ instead of $\sigma\sqrt{n}$. Another way to reach the same result is to study

$$\frac{W_n - \lambda}{\sigma_{W_n}}; \tag{2.5}$$

Date: February 9, 2005.

here W_n is normalized so it has mean λ, and of course its standard deviation is σ_{W_n}. Substituting gives

$$\frac{W_n - \lambda}{\sigma_{W_n}} \ = \ \frac{\frac{1}{n}Y_n - \lambda}{\sigma/n} \ = \ \frac{Y_n - n\lambda}{\sigma\sqrt{n}} \ = \ Z_n. \tag{2.6}$$

Thus, we obtain the same quantity as before. We now use

$$M_{\frac{X+a}{b}}(t) \ = \ e^{at/b} M_X(t/b) \tag{2.7}$$

and the moment generating function of a sum of independent variables is the product of the moment generating functions to find the moment generating function of Z_n. Note $\frac{Y_n - n\lambda}{\sigma\sqrt{n}} = \sum_i \frac{X_i - \lambda}{\sigma\sqrt{n}}$. Therefore

$$\begin{aligned} M_{Z_n}(t) \ &= \ M_{\frac{Y_n - n\lambda}{\sigma\sqrt{n}}}(t) \\ &= \ M_{\sum_i \frac{X_i - \lambda}{\sigma\sqrt{n}}}(t) \\ &= \ \prod_i M_{\frac{X_i - \lambda}{\sigma\sqrt{n}}}(t) \\ &= \ \prod_i e^{\frac{-\lambda t}{\sigma\sqrt{n}}} \ M_X\left(\frac{t}{\sigma\sqrt{n}}\right) \\ &= \ \prod_i e^{\frac{-\lambda t}{\sigma\sqrt{n}}} \ e^{\lambda\left(e^{\frac{t}{\sigma\sqrt{n}}} - 1\right)}. \end{aligned} \tag{2.8}$$

We now Taylor expand the exponential, using

$$e^u \ = \ \sum_{k=0}^{\infty} \frac{u^k}{k!} \ = \ 1 + u + \frac{u^2}{2} + \frac{u^3}{6} + \cdots. \tag{2.9}$$

This is one of the most important Taylor expansion we will encounter. Thus the exponential in (2.8) is

$$e^{\frac{t}{\sigma\sqrt{n}}} \ = \ 1 + \frac{t}{\sigma\sqrt{n}} + \frac{t^2}{2\sigma^2 n} + \frac{t^3}{6\sigma^3 n\sqrt{n}} + \cdots. \tag{2.10}$$

The important thing to note is that after subtracting 1, the first piece is $\frac{t}{\sigma\sqrt{n}}$, the next piece is $\frac{t^2}{2\sigma^2 n}$, and the remaining pieces are dominated by a geometric series (starting with the cubed term) with $r = \frac{t}{\sigma\sqrt{n}}$. Thus, the contribution from all the other terms is of size at most some constant times $\frac{t^3}{n\sqrt{n}}$. For large n, this will be negligible, and we write errors like this as $O\left(\frac{t^3}{n\sqrt{n}}\right)$.

Thus, (2.8) becomes

$$\begin{aligned} M_{Z_n}(t) \ &= \ \prod_i e^{\frac{-\lambda t}{\sigma\sqrt{n}}} \ e^{\lambda\cdot\left(\frac{t}{\sigma\sqrt{n}} + \frac{t^2}{2\sigma^2 n} + O\left(\frac{t^3}{n\sqrt{n}}\right)\right)} \\ &= \ \prod_i e^{\frac{\lambda t^2}{\sigma^2 n} + O\left(\frac{t^3}{n\sqrt{n}}\right)} \\ &= \ e^{\frac{t^2}{2} + O\left(\frac{t^3}{\sqrt{n}}\right)} \end{aligned} \tag{2.11}$$

where the last line follows from the fact that we have a product over n identical terms, and as the variance of X is $\sigma^2 = \lambda$ (for X Poisson with parameter λ), we see $\frac{\lambda}{\sigma^2} = 1$. Thus, for all t, as $n \to \infty$ the moment generating function of Z_n tends to $e^{\frac{t^2}{2}}$, which is the moment generating function of the standard normal. This completes the proof. □

Remark 2.1. We only need to Taylor expand far enough to get the main term (which has a finite limit as $n \to \infty$) and then estimate the size of the error term (which tends to zero as $n \to \infty$).