

262 Chapter 8 Estimation of Parameters and Fitting of Probability Distributions

A standard statistical technique for addressing this question is to derive the sampling distribution of the estimate or an approximation to that distribution. The statistical model stipulates that the individual counts X_i are independent Poisson random variables with parameter λ_0 . Letting $S = \sum X_i$, the parameter estimate $\hat{\lambda} = S/n$ is a random variable, the distribution of which is called its sampling distribution. Now from Example E in Section 4.5, the distribution of the sum of independent Poisson random variables is Poisson distributed, so the distribution of S is Poisson ($n\lambda_0$). Thus the probability mass function of $\hat{\lambda}$ is

$$\begin{aligned} P(\hat{\lambda} = v) &= P(S = nv) \\ &= \frac{(n\lambda_0)^{nv} e^{-n\lambda_0}}{(nv)!} \end{aligned}$$

for v such that nv is a nonnegative integer.

Since S is Poisson, its mean and variance are both $n\lambda_0$, so

$$\begin{aligned} E(\hat{\lambda}) &= \frac{1}{n} E(S) = \lambda_0 \\ \text{Var}(\hat{\lambda}) &= \frac{1}{n^2} \text{Var}(S) = \frac{\lambda_0}{n} \end{aligned}$$

From Example A in Section 5.3, if $n\lambda_0$ is large, the distribution of S is approximately normal; hence, that of $\hat{\lambda}$ is approximately normal as well, with mean and variance given above. Because $E(\hat{\lambda}) = \lambda_0$, we say that the estimate is **unbiased**: the sampling distribution is centered at λ_0 . The second equation shows that the sampling distribution becomes more concentrated about λ_0 as n increases. The standard deviation of this distribution is called the **standard error** of $\hat{\lambda}$ and is

$$\sigma_{\hat{\lambda}} = \sqrt{\frac{\lambda_0}{n}}$$

Of course, we can't know the sampling distribution or the standard error of $\hat{\lambda}$ because they depend on λ_0 which is unknown. However, we can derive an approximation by substituting $\hat{\lambda}$ for λ_0 and use it to assess the variability of our estimate. In particular, we can calculate the **estimated standard error** of $\hat{\lambda}$ as

$$s_{\hat{\lambda}} = \sqrt{\frac{\hat{\lambda}}{n}}$$

For this example, we find

$$s_{\hat{\lambda}} = \sqrt{\frac{24.9}{23}} = 1.04$$

At the end of this section, we will present a justification for using $\hat{\lambda}$ in place of λ_0 .

In summary, we have found that the sampling distribution of $\hat{\lambda}$ is approximately normal, centered at the true value λ_0 with standard deviation 1.04. This gives us a reasonable assessment of the variability of our parameter estimate. For example, because a normally distributed random variable is unlikely to be more than two standard deviations away from its mean, the error in our estimate of λ is unlikely to be more than 2.08. We thus have not only an estimate of λ_0 , but also an understanding of the inherent variability of that estimate.