Tutorial # 9

(~ 21.1.14)

Distribute randomly n points over interval $(0, t)$. (t — fixed number)

Let $\Delta t \subset (0,t)$, $p = \frac{\Delta t}{t}$ — probability of finding a particle in Δt

Then, prob. of finding m particles in Δt

$$p(m) = \frac{n!}{m!(n-m)!} \cdot p^m (1-p)^{n-m} = \frac{n!}{m!(n-m)!} \left(\frac{\Delta t}{t}\right)^m \cdot \left(1 - \frac{\Delta t}{t}\right)^{n-m}$$

Binomial distribution

Note: $\sum_{m=0}^{n} m\, p(m) = \ldots = np = \underbrace{\frac{n \Delta t}{t}}_{=: \lambda \Delta t} =: \mu$

average value

Let $\Delta t \to 0$ (or $n \to \infty$) s.t. $\mu = \lambda \Delta t = \text{const}$ (or, equivalently, $n\Delta t = \text{const} = \mu t$)

Then, $p(m) = \frac{n!}{(n-m)!\, m!} \cdot \frac{(\lambda \Delta t)^m}{n^m} \cdot \left(1 - \frac{\lambda \Delta t}{n}\right)^{n-m} =$

$$= \underbrace{\frac{n!}{(n-m)!\, n^m}}_{} \cdot \frac{(\lambda \Delta t)^m}{m!} \cdot \underbrace{\left(1 - \frac{\lambda \Delta t}{n}\right)^n}_{\xrightarrow[n\to\infty]{} e^{-\lambda \Delta t}} \cdot \underbrace{\frac{1}{\left(1 - \frac{\lambda \Delta t}{n}\right)^m}}_{\xrightarrow[n\to\infty]{} 1} \longrightarrow \frac{(\lambda \Delta t)^m}{m!} e^{-\lambda \Delta t}$$

Poisson distribution

n factors

$$= \frac{n(n-1)(n-2)\cdots(n-m+1)}{n^m} = \underbrace{\left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right)}_{\xrightarrow[n\to\infty]{} 1} \longrightarrow 1$$

We can also look at Poisson distribution as that arising from a counting process:

Let $X(t)$ be the # of events occured in the time t starting from $t=0$.

Divide $[0,t]$ into n subintervals of length $\Delta t = \frac{t}{n}$

Assume:

1/ at each Δt there's at most 1 event (i.e. take Δt small enough)

2/ probability p that an event occurs in a Δt subinterval is $\lambda \Delta t$, $\lambda > 0$.

3) occurences in subintervals are independent

Then, by the binomial distribution:

$$P\big(X(\overset{n\Delta t}{t}) = k\big) = \frac{n!}{k!(n-k)!}\, p^k (1-p)^{n-k} =$$

= prob. that k events happen in time $t = n\Delta t$

= prob. that k subintervals out of n are not empty

$$= \frac{n!}{k!(n-k)!} (\lambda \Delta t)^k (1-\lambda \Delta t)^{n-k} = \frac{n!}{k!(n-k)!} \frac{(\lambda t)^k}{n^k} \cdot \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

$$\xrightarrow[\text{as } n\to\infty]{} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad \text{(as before)}$$

Poisson distribution

N 22.1.3

Exponential distribution:

$p(x;\lambda) = \lambda e^{-\lambda x}, \ x > 0$ — assume form of distribution is known for population

λ — parameter to estimate

Sample: $x_1, \dots, x_n$ — observation of independent discrete rand. variables $X_1, \dots, X_n$

Introduce the likelihood function:

$$L(x_1, \dots, x_n; \lambda) = p(x_1, \lambda) \cdot \ldots \cdot p(x_n, \lambda)$$

(prob. that $X_1 = x_1, \dots, X_n = x_n$ are observed assuming λ is the true value of population parameter.

Since particular realization was observed, it makes sense that, at the true value of λ, the prob. L is maximal, i.e.

$$\underbrace{\frac{\partial L}{\partial \lambda} = 0, \quad \frac{\partial^2 L}{\partial \lambda^2} < 0.}_{\text{necess. cond. of max assuming differentiability of } L}$$

note:

$$\frac{\partial L}{\partial \lambda} = 0 \iff \overbrace{\frac{\partial \log L}{\partial \lambda}}^{= \frac{1}{L} \cdot \frac{\partial L}{\partial \lambda}} = 0$$

$$\frac{\partial^2 L}{\partial \lambda^2} < 0 \iff \underbrace{\frac{\partial^2 \log L}{\partial \lambda^2}}_{= \frac{1}{L} \frac{\partial^2 L}{\partial \lambda^2} - \frac{1}{L^2} \underbrace{\left(\frac{\partial L}{\partial \lambda}\right)^2}_{= 0}} < 0$$

Because of multiple products and exponential, it's more convenient to work with $\log L$ rather than L itself.

(~ 21.1.16)

We count # of the red blood cells in a small volume of blood under microscope.

$\mu = $ (Average count of a healthy individual) $= 8$

a) $\underline{P}(k=4) = \frac{\mu^4}{4!} \cdot e^{-\mu} = \frac{8^4 \cdot e^{-8}}{4 \cdot 3 \cdot 2} = \frac{2^9 \cdot e^{-8}}{3}$

prob. that the count of a healthy individual contains only 4 red blood cells

(—"— 4 or fewer

b) $\underline{P}(k \leq 4) = \sum_{k=0}^{4} \underline{P}(k) = e^{-8} \cdot \left(\frac{2^9}{3} + \frac{2^8}{3} + 2^5 + 2^3 + 1 \right)$

$$\log L = \log\lambda - \lambda x_1 + \log\lambda - \lambda x_2 + \ldots = n\log\lambda - \lambda\sum_{k=1}^{n} x_k$$

$$\frac{\partial \log L}{\partial\lambda} = 0 \Rightarrow \frac{n}{\lambda} - \sum_{k=1}^{n} x_k = 0 \Rightarrow \lambda = \frac{n}{\sum_{k=1}^{n} x_k}$$

$$\Rightarrow \underbrace{\hat{\lambda}}_{\text{estimate of } \lambda} = \frac{1}{\bar{x}}, \text{ where } \bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \text{ - sample mean}$$

$$\left\{ \frac{\partial^2 \log L}{\partial\lambda^2} = -\frac{n}{\lambda^2} < 0 \Rightarrow \hat{\lambda} = \frac{1}{\bar{x}} \text{ is indeed maximum.} \right.$$

№ 22.1.4

$$p(K, p_0) = \binom{\overset{\text{fixed}}{N}}{K} \cdot \underbrace{p_0}_{\text{to estimate}}^{K} \cdot (1-p_0)^{N-K}$$

$$L(K_1, \ldots, K_n; p_0) = \binom{N}{K_1}\cdots\binom{N}{K_n} \cdot p_0^{\sum_{i=1}^{n} K_i} \cdot (1-p_0)^{N\cdot n - \sum_{i=1}^{n} K_i}$$

$$\frac{\partial \log L}{\partial p_0} = 0 \Rightarrow \frac{\sum_{i=1}^{n} K_i}{p_0} - \frac{N n - \sum_{i=1}^{n} K_i}{1-p_0} = 0$$

$$\Rightarrow \sum_{i=1}^{n} K_i \cdot (1-p_0) - N\cdot n\cdot p_0 + \cancel{\sum_{i=1}^{n} K_i \cdot p_0} = 0 \Rightarrow \hat{p}_0 = \frac{\sum_{i=1}^{n} K_i}{N\cdot n} = \frac{\bar{K}}{N}$$

$$\left\{ \frac{\partial^2 \log L}{\partial p_0^2} = -\frac{\sum_{i=1}^{n} K_i}{p_0^2} - \frac{\overbrace{N\cdot n - \sum_{i=1}^{n} K_i}^{\geq 0}}{(1-p_0)^2} < 0 \Rightarrow \hat{p}_0 \text{ is maximum} \right.$$