Lecture 23 - 10/14

Friday, October 14, 2022 10:05 AM

Review:

Guassian distribution

- X has guassian (normal) distribution, $X \sim N(\mu, \sigma^2)$ if

$$f_X(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right), \quad u \in \mathbb{R}$$

μ: mean σ^2: variance

- Standardized version of X, $\tilde{X} = \frac{X-\mu}{\sigma}$ is distributed as $N(0,1)$.

- $$F_{\tilde{X}}(a) = P(\tilde{X} \leq a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{u^2}{2}\right) du = \Phi(a)$$

- $$F^c_{\tilde{X}}(a) = P(\tilde{X} > a) = \int_{a}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = 1 - \Phi(a) = Q(a)$$

- Notice that $Q(a) = \Phi(-a)$ because of symmetry!

- We can write:

$$\begin{aligned} P(a < X \leq b) &= P(X \leq b) - P(X \leq a) \\ &= P\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{a}\right) - \Phi\left(\frac{a-\mu}{a}\right) \end{aligned}$$

$$P(X > b) = P\left(\frac{X-\mu}{\sigma} > \frac{b-\mu}{\sigma}\right) = Q\left(\frac{b-\mu}{\sigma}\right) = \Phi\left(-\frac{b-\mu}{\sigma}\right)$$

Today:

1. The central limit theorem and approximation
2. ML parameter estimate for continuous-typ variables

---

① The central limit theorem and approximation

Loosely speaking: if many independent random variables are added together, and if each of them are small in magnitude compared to sum, the the sum has approximately normal distribution.

• If $X_i \sim Ber(p)$ and independent of each other, then $S_n = X_1 + \dots + X_n$ is $Bi(n,p)$. We can approximat S_n by a guassian random variable:

$$P(S_n \leq k) \approx P(W \leq k)$$

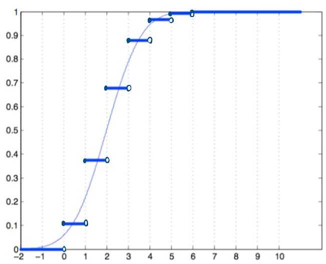
where $W \sim N(\mu, \sigma^2)$, $\mu = E[S_n] = np$, $\sigma^2 = Var(S_n) = np(1-p)$. Approximation is mostly accurate if

(i) $\mu = np$ is moderately large

(ii) $\sigma = np(1-p)$ is moderately large

(iii) The probability of the event of interest is not too small.

• If we plot the cdf of S_n and W: $n=10$, $p=0.2$. $\mu = np = 2$, $\sigma = \sqrt{np(1-p)} = 1.6$



cdf of S_n: line with jumps

cdf of W: continuous line

Observations:

(i) cdf of S_n is piecewis constant, cdf of W is continuous: they cannot be close everywhere.

(ii) suprisingly, they are close at midle points:

$$P(S_n \leq 2) = P(S_n \leq 2.5) \approx P(W \leq 2.5)$$

• Guassian approximation with the continuity correction:

$$P(S_n \leq k) \approx P(W \leq k + 0.5)$$

$$P(S_n > k) = 1 - P(S_n \leq k) \approx 1 - P(W \leq k + 0.5) = P(W \geq k + 0.5)$$

$$P(S_n \geq k) \approx P(W > k - 0.5)$$

where $S_n \sim Bi(n,p)$ and $W \sim N(\mu, \sigma^2)$, $\mu = np$, $\sigma = \sqrt{np(1-p)}$. In particular:

$$P(S_n = k) = P(S_n \leq k) - P(S_n \leq k-1)$$

$$\approx P(W \leq k + 0.5) - P(W \leq k - 0.5) = \int_{k-0.5}^{k+0.5} f_W(u)\, du = \Phi\left(\frac{k+0.5-\mu}{\sigma}\right) - \Phi\left(\frac{k-0.5-\mu}{\sigma}\right)$$

.CLT for Binomial random variable

Theorem (DeMoivre–Laplace limit theorem)

Suppose that $S_{n,p} \sim Bi(n,p)$, $p \in (0,1)$ fixed.

$$\lim_{n \to \infty} P\left\{ \frac{S_{n,p} - np}{\sqrt{np(1-p)}} \leq c \right\} = \Phi(c)$$

i.e., for large n, $S_{n,p}$ has the same cdf as a guassian random variable.

.What happened to continuity correction?

$$P\left\{ \frac{S_{n,p} - np}{\sqrt{np(1-p)}} \leq c \right\} \approx P\left\{ \frac{X - 0.5 - np}{\sqrt{np(1-p)}} \leq c \right\}$$ with continuity correction

$$\approx P\left\{ \frac{X - np}{\sqrt{np(1-p)}} \leq c \right\}$$ n is large so that $np(1-p)$ is large

---

3.19. **[Heads minus tails]**

Suppose a fair coin is flipped 100 times, and A is the event:
$A = \{\ |(\text{number of heads}) - (\text{number of tails})| \geq 10\ \}$.

(a) Let S denote the number of heads. Express A in terms of S. Specifically, identify which values of S make A true.

(b) Using the Gaussian approximation with the continuity correction, express the approximate value of $P(A)$ in terms of the Q function.

Solution:

(a) Notice that # of tails = $100 - S$. Hence

$$A = \{ |\text{number of heads} - \text{number of tails}| \geq 10 \} = \{ |S - (100 - S)| \geq 10 \}$$
$$= \{ |S - 50| \geq 5 \}$$
$$= \{ S \geq 55 \} \cup \{ S \leq 45 \}$$

(b) $P(\{S \geq 55\} \cup \{S \leq 45\}) = P(S \geq 55) + P(S \leq 45)$.

Notice that $P(S \leq 45) \approx P(W \leq 45 + 0.5)$ and $P(S \geq 55) \approx P(W \geq 55 - 0.5)$

where $W \sim N(\mu, \sigma^2)$, $\mu = np = 50$, $\sigma = \sqrt{np(1-p)} = 5$. Hence,

$$P(\{S \geq 55\} \cup \{S \leq 45\}) \approx P(W \leq 45.5) + P(W \geq 54.5)$$

$$= \Phi\left(\frac{45.5 - 5}{5}\right) + \Phi\left(\frac{54.5 - 5}{5}\right)$$

⊛ ML parameter estimate for continuous-typ variables

Suppose that the random variable X is continuous type and its pdf belongs to a familly of distributions $f_\theta(\cdot)$ (example: X is exponential random variable, in here θ is the parameter of pdf, i.e., $X \sim \exp(\theta)$)

Suppose that we observe $X = u$. Notice that $P(X=u) = 0$ since X is continuous type. Recall that for $\varepsilon > 0$ small enough.

$$P(u-\varepsilon < X < u+\varepsilon) \approx 2\varepsilon \cdot f_\theta(u)$$

Maximum likelihood estimation for continuous type suggests maximizing $f_\theta(u)$ over θ, given the observation is $X=u$.

$$\hat{\theta}_{ML} = \underset{\theta > 0}{\arg\max}\, f_\theta(u)$$

3.23. **[ML parameter estimation of the rate of a Poisson process]**
Calls arrive in a call center according to a Poisson process with arrival rate λ (calls/minute). Derive the maximum likelihood estimate $\widehat{\lambda}_{ML}$ if k calls are received in a t minute interval.

We consider two different interpretation of the problem.

(a) k^{th} call is recieved at time t.

Notice that the time of kth arrival is T_k, which has Erlang distribution with parameters (k, λ).

$$\hat{\lambda}_{ML} = \underset{\lambda > 0}{\arg\max}\, f_{T_k}(t) = \underset{\lambda > 0}{\arg\max}\, \frac{\lambda e^{-\lambda t} (\lambda t)^{k-1}}{(k-1)!}$$

To solve above, we need to solve

$$\left.\frac{d f_{T_k}(t)}{d\lambda}\right|_{\lambda = \hat{\lambda}_{ML}} = 0 \Rightarrow \left.\frac{d}{d\lambda}\left(\lambda e^{-\lambda t} \lambda^{k-1}\right)\right|_{\lambda = \hat{\lambda}_{ML}} = 0.$$

(b) In the first t minutes we recieved k calls, i.e., $N_t = k$.

Notice that $N_t \sim \mathrm{Poi}(\lambda t)$

$$\hat{\lambda}_{ML} = \underset{\lambda > 0}{\arg\max}\, P(N_t = k) = \underset{\lambda > 0}{\arg\max}\, \frac{e^{-\lambda t} (\lambda t)^k}{k!}$$

$$\hat{\lambda}_{ML} = \underset{\lambda>0}{\operatorname{argmax}} \, P(N_t = k) = \underset{\lambda>0}{\operatorname{argmax}} \, \frac{e^{-\lambda t}(\lambda t)}{k!}$$

Hence, we need to solve

$$\left.\frac{d\,P(N_t=k)}{d\lambda}\right|_{\lambda=\hat{\lambda}_{ML}} = 0 \Rightarrow \left.\frac{d(e^{-\lambda t}\lambda^k)}{d\lambda}\right|_{\lambda=\hat{\lambda}_{ML}} = 0$$

Interestingly, they both gives the same λ. Can you interpret why?