Lecture 21

3.5.3 Erlang distribution

Bernoulli Process: # of trials to get r successes ⇒ Neg. Binomial

$\Downarrow$

Limit of scaled Bernoulli Process
T_r: time to get the r successes ~ Erlang Distribution

$$\{T_r > t\} = \{r-1 \text{ or less successes by time } t\}$$
$$= \{N_t \le r-1\}$$

$$F_{T_r}(t) = 1 - P(N_t \le r-1) = 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

$$f_{T_r}(t) = \frac{dF_{T_r}(t)}{dt} = -\sum_{k=0}^{r-1}\left(\frac{-\lambda e^{-\lambda t}(\lambda t)^k}{k!} - \frac{e^{-\lambda t}\lambda^k k\, t^{k-1}}{k!}\right)$$

$$= e^{-\lambda t}\left(\sum_{k=0}^{r-1} \lambda \frac{(\lambda t)^k}{k!} - \sum_{k=1}^{r-1} \frac{\lambda^k t^{k-1}}{(k-1)!}\right)$$

$$= e^{-\lambda t}\left(\sum_{k=0}^{r-1} \lambda \frac{(\lambda t)^k}{k!} - \sum_{k'=0}^{r-2} \frac{\lambda^{k'+1} t^{k'}}{k'!}\right) = e^{-\lambda t}\left(\frac{\lambda^r t^{r-1}}{(r-1)!}\right)$$

Def) A r.v. T_r is an Erlang r.v. with parameters r and λ if the pdf is given by

$$f_{T_r}(t) = \begin{cases} \dfrac{e^{-\lambda t}\lambda^r t^{r-1}}{(r-1)!} & \text{if } t \ge 0 \\ 0 & \text{o/w.} \end{cases}$$

Exercise: show $E[T_r] = \frac{r}{\lambda}$ and $Var(T_r) = \frac{r}{\lambda^2}$

3.6.1 Linear scaling of pdf.

For simplicity, it is useful to covert a random variable with mean μ and variance σ^2 to a same-kind random variable with mean 0 and variance 1.

How? $\Rightarrow$ Using a linear scaling.

Let a r.v. X has pdf $f_X(u)$.

Let $Y = aX + b$ $\quad (a > 0)$

$$\Rightarrow f_Y(u) = f_X\left(\frac{u-b}{a}\right)\frac{1}{a}$$

$$F_Y(u) = P(aX+b \le u) = P\left(X \le \frac{u-b}{a}\right) = F_X\left(\frac{u-b}{a}\right)$$

$$f_Y(u) = \frac{d}{du}F_Y(u) = \frac{d}{du}F_X\left(\frac{u-b}{a}\right) = f_X\left(\frac{u-b}{a}\right)\cdot\frac{1}{a}$$

$$\Rightarrow E[Y] = aE[X] + b, \quad Var(Y) = a^2 Var(X)$$

Ex) For a given r.v. X with mean μ and variance σ^2, how to choose a and b to make a standardized random variable Y of X?

※ Standardized r.v. : the same kind r.v. with mean 0 and variance 1.

$$1 = a^2 Var(X) \Rightarrow a = \frac{1}{\sigma}$$

$$E[Y] = aE[X] + b = 0 \Rightarrow b = -\frac{\mu}{\sigma}$$

3.6.2. Gaussian (normal) Random Variable

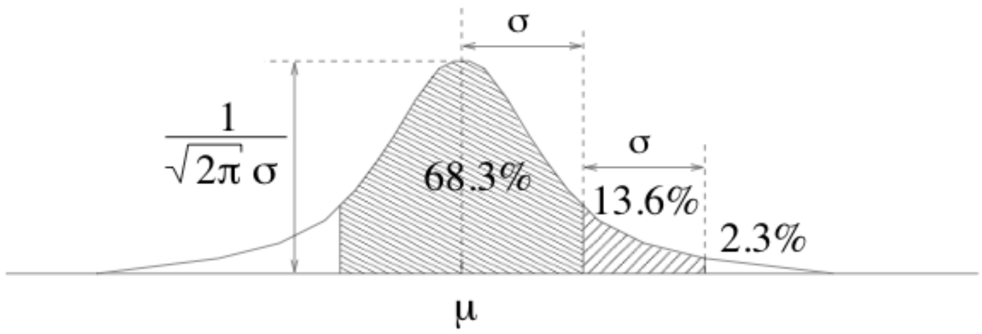
Def) A r.v. X is a gaussian (also called normal) if the pdf is

$$f_X(u) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(u-\mu)^2}{2\sigma^2}}$$

where the parameters μ and σ^2.

Often denoted by $N(\mu, \sigma^2)$.

* Sketch of pdf



* Mean and variance of X

$$E[X] = \mu, \quad Var(X) = \sigma^2.$$

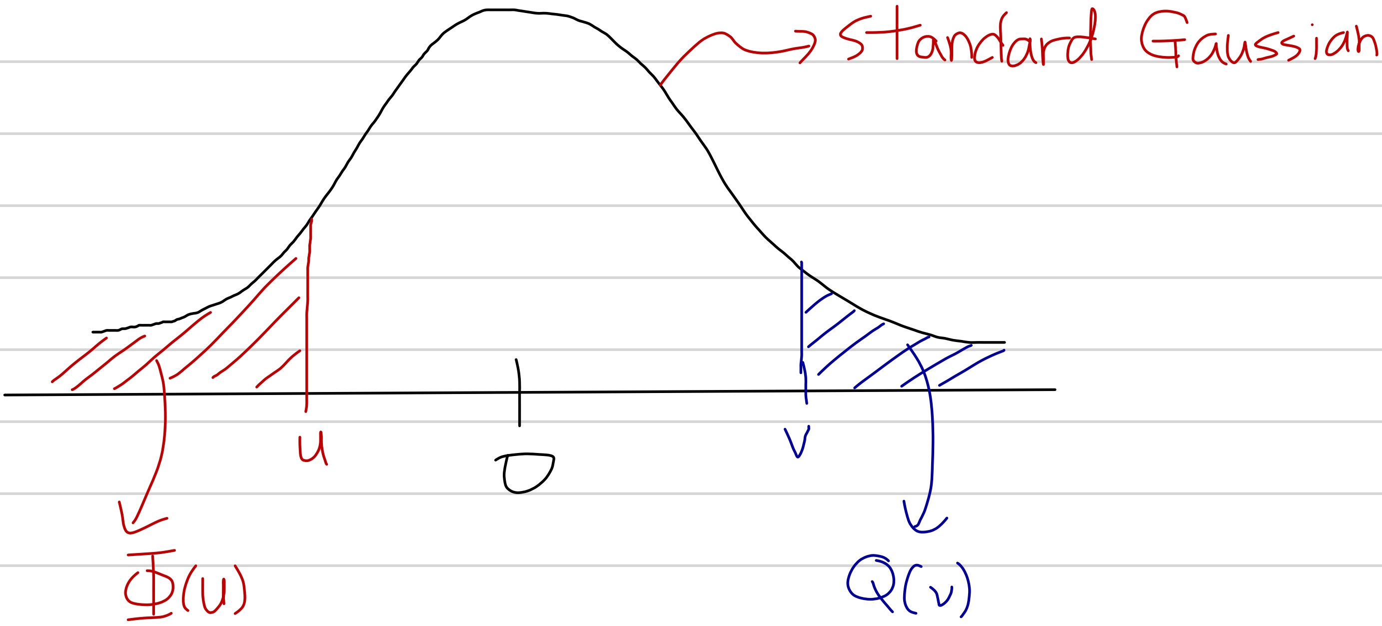
Why? Exercise (refer to the lecture note by Prof. Hajek)

Standardized normal distribution $\sim N(0,1)$

Φ (Phi) function & Q function

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{u^2}{2}\right) du$$

$$Q(u) = \int_{u}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{u^2}{2}\right) du.$$



How to compute such prob. in $N(\mu, \sigma^2)$?

$$N(\mu, \sigma) \xrightarrow[\text{linear scaling}]{} N(0,1) \longrightarrow \text{use } Q, \Phi$$

Ex) $X \sim N(3,4)$

standard Gaussian   2

Q1) $P(X \geq 7) = P\left(\frac{X-3}{2} \geq \frac{7-3}{2}\right) = Q(2)$

Q2) $P\{(X-1)(X-7) \geq 0\} = P(X \leq 1 \text{ or } X \geq 7\} = P(X \leq 1) + P(X \geq 7)$

$$= P\left(\frac{X-3}{2} \leq -1\right) + P\left(\frac{X-3}{2} \geq 2\right)$$

$$= \Phi(-1) + Q(2)$$

Lecture 22

3.6.3 The central limit theorem.

$X_1, X_2, \cdots, X_n$: independent r.v.s with mean μ and variance σ^2

$$X = \sum_{k=1}^{n} X_k \quad (\text{mean } n\mu, \text{ variance } n\sigma^2)$$

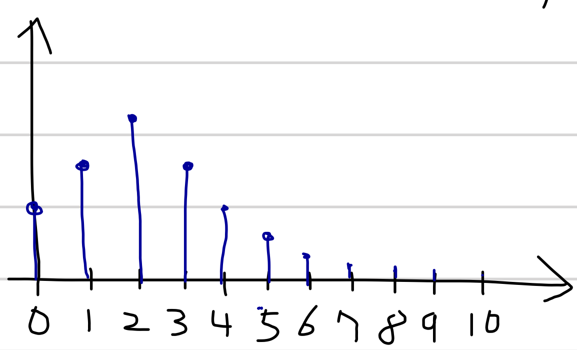
$Y \sim N(0,1)$

For large n, $P\left\{ \frac{X - n\mu}{\sqrt{n\sigma^2}} \le c \right\} \approx P\{Y \ge c\}$

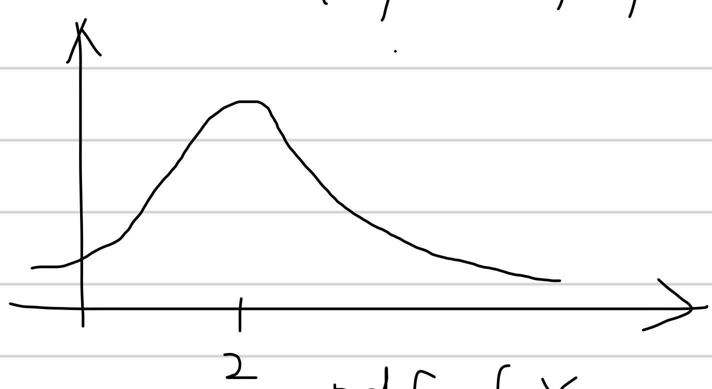
In other words, let $Y \sim N(n\mu, n\sigma^2)$

$$P(S_n \ge c) \approx P(Y \ge c)$$

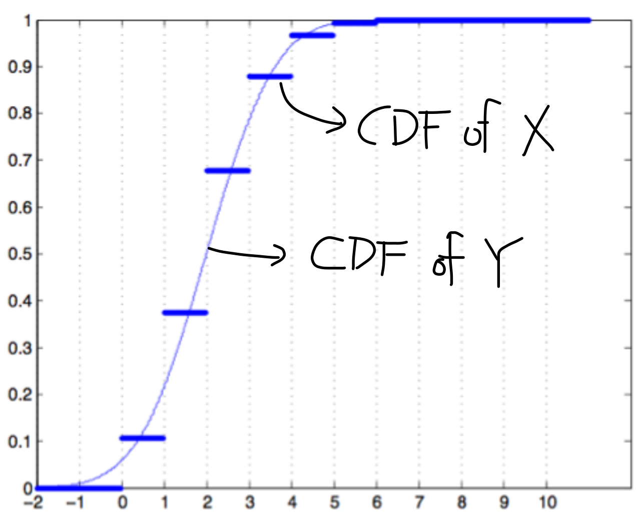
Ex) $X \sim \text{Binomial}(n=10, p=0.2)$ VS. $Y \sim N(np=2, np(1-p)=1.6)$



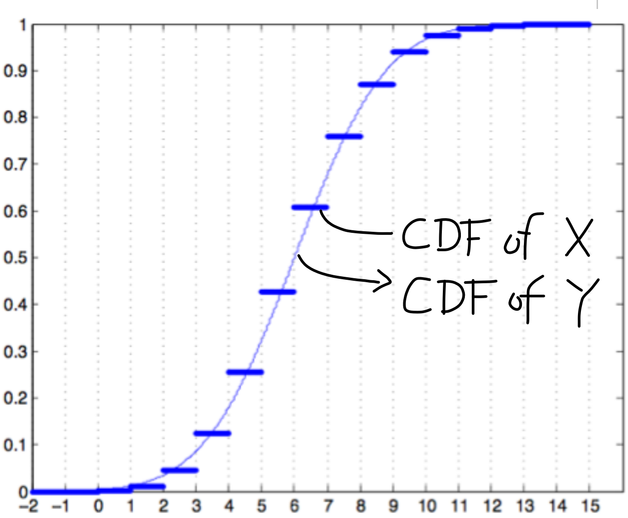
pmf of X



pdf of Y



$n = 10.$



$n = 30$

Gaussian approximation with the continuity correction.

X is discrete-type

$\Rightarrow P(X \le 2) = P(X \le 2.5) \approx P(Y \le 2.5)$

For large n, $P(X \le 2.5)$ is a good approximation of $P(X \le 2)$

$$\Rightarrow \quad P\{X \le k\} \approx P(Y \le k+0.5)$$
$$P\{X \ge k\} \approx P(Y \ge k-0.5)$$

"De Moivre-Laplace limit theorem" Suppose $S_{n,p}$ is a binomial random variable with parameters (n,p). For fixed p,

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

Ex) Toss a fair coin 1000 times.

$X = \#$ of heads $\sim$ Binomial$(1000, \frac{1}{2})$

Q1) Find $P\{X \ge K\} \approx 0.01$

$$\Rightarrow P\{X \ge K\} = P\{X \ge K-0.5\} = P\left\{\frac{X-\mu}{\sigma} \ge \frac{K-0.5-\mu}{\sigma}\right\}$$
$$\approx Q\left(\frac{K-0.5-\mu}{\sigma}\right) = 0.01$$

$\mu = 500$, $\sigma = 15.8$, $Q(2.325) = 0.01$

$\Rightarrow K = \mu + 2.325 \times \sigma + 0.5 = 537.26$.

$\Rightarrow K$ must be 537 or 538

ML parameter estimation for continuous r.v.

For a r.v. X with $f_\theta(u) \to$ pdf of X for a given parameter θ

$\Rightarrow$ $X = j$ observed. How to estimate θ?

Prob. of observing $X \in [j-\frac{\varepsilon}{2}, j+\frac{\varepsilon}{2}]$

$$P[X \in [j-\tfrac{\varepsilon}{2}, j+\tfrac{\varepsilon}{2}]] \approx f_\theta(j) \cdot \varepsilon$$

Thus $\hat{\theta}_{ML}(u) = \arg\max_\theta f_\theta(j)\varepsilon = \arg\max f_\theta(j)$

Ex) $X \sim N(5, \sigma^2)$. σ is unknown

Given $X = u$ observed, $\hat{\sigma}_{ML}$?

$$f_X(u) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-5)^2}{2\sigma^2}} \Rightarrow \ln f_X = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln(\sigma^2) - \frac{(u-5)^2}{2\sigma^2}$$

$$\frac{d}{d\sigma^2}(\ln f_X) = -\frac{1}{2\sigma^2} + \frac{(u-5)^2}{2(\sigma^2)^2} = \frac{1}{2\sigma^2}\left(\frac{(u-5)^2}{\sigma^2} - 1\right)$$

$$\begin{cases} >0 \text{ if } \sigma^2 < (u-5)^2 \\ <0 \text{ if } \sigma^2 > (u-5)^2 \end{cases}$$

$\Rightarrow$ Maximum of $f_X(u)$ is attained at $\sigma = (u-5)$

$\Rightarrow \hat{\sigma}_{ML} = (u-5)$