EXERCISES

(1) Let $X_1 \ldots X_n$ be a simple sample from the distribution given by the density $f_\theta(x) = \begin{cases} e^{-(x-\theta)} & x \geq \theta \\ 0 & x < \theta \end{cases}$ $\theta \in \mathbb{R}$ unknown. Find the MLE.

- If $\theta = 0$, we have the exponential distribution of parameter 1: $X \rightsquigarrow \mathcal{E}(1)$

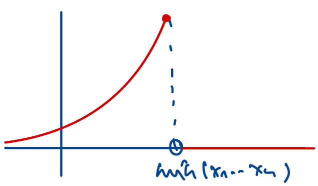
> Note: The exponential distribution of parameter $\beta > 0$ $\mathcal{E}(\beta)$
>
> has density $f(x) = \begin{cases} 0 & x < 0 \\ \beta e^{-\beta x} & x \geq 0 \end{cases}$

We will maximize the likelihood function

$$L_{X_1 \ldots X_n}(\theta) = \prod_{i=1}^{n} f_\theta(x_i) = \begin{cases} 0 & \exists i = 1 \ldots n \,/\, x_i < \theta \\ \prod_{i=1}^{n} e^{-(x_i-\theta)} = e^{n\theta - \sum_{i=1}^{n} x_i} & x_i \geq 0 \ \forall i = 1 \ldots n \end{cases} = \begin{cases} 0 & \theta > \min(x_1 \ldots x_n) \\ e^{n\theta - \sum_{i=1}^{n} x_i} & \theta \leq \min(x_1 \ldots x_n) \end{cases}$$

independent observations

$L'_{X_1 \ldots X_n}(\theta) = n e^{n\theta - \sum_{i=1}^{n} x_i} \neq 0$. However, we can see the maximum value is obtained at $\theta = \min(x_1 \ldots x_n)$



$\hat{\theta} = \hat{\theta}(x_1 \ldots x_n) = \min(x_1 \ldots x_n)$ is the MLE.

② Let $X_1 \dots X_n$ be a simple sample from $N(\mu, \gamma^2 \mu^2)$, $\mu \in \mathbb{R}$ unknown, $\gamma \in \mathbb{R}^+$ known. Find the estimator $\hat{\mu}$ on the parameter μ of the form $\hat{\mu} = \sum_{i=1}^{n} c_i X_i$ / $\forall \mu \in \mathbb{R}$, $E_\mu[(\hat{\mu}-\mu)^2]$ is the smallest possible.

Mean square error of the estimator $\hat{\mu}$

$E[\hat{\mu}] = \mu$ var $\hat{\mu} \to$ min $\to 0$

$$Var[\hat{\mu}-\mu] = E[(\hat{\mu}-\mu)^2] + E[\hat{\mu}-\mu]^2 = E[(\hat{\mu}-\mu)^2]$$

$$\min \Leftrightarrow E_\mu[(\hat{\mu}-\mu)^2] = Var_\mu[\hat{\mu}-\mu] + E_\mu[\hat{\mu}-\mu]^2 = Var_\mu[\hat{\mu}] + (E_\mu[\hat{\mu}]-\mu)^2 =$$

$$Var_\mu[\sum_{i=1}^{n} c_i X_i] + (E_\mu[\sum_{i=1}^{n} c_i X_i] - \mu)^2 = \sum_{i=1}^{n} c_i^2 Var_\mu X_i + \left(\sum_{i=1}^{n} c_i E_\mu[X_i] - \mu\right)^2 =$$

$$\mu^2 \left(\gamma^2 \sum_{i=1}^{n} c_i^2 + \left(\sum_{i=1}^{n} c_i - 1\right)^2\right)$$

We need to minimize $f(c_1 \dots c_n) = \left(\gamma^2 \sum_{i=1}^{n} c_i^2 + \left(\sum_{i=1}^{n} c_i - 1\right)^2\right)$. because due to the form of the expression, it is minimized $\forall \mu \in \mathbb{R}$.

$k = 1 \dots n$ $\quad \frac{\partial f}{\partial c_k}(c_1 \dots c_n) = 2\gamma^2 c_k + 2\left(\sum_{i=1}^{n} c_i - 1\right) = 0$

We know the function is convex and tends to $+\infty$ $\Longrightarrow$ its local minimum is also global.

$\gamma^2 c_k = 1 - \sum_{i=1}^{n} c_i \Rightarrow c_1 = \dots = c_n$

$\gamma^2 c_k = 1 - n c_k \Rightarrow c_k = \frac{1}{n+\gamma^2} \quad \forall k = 1 \dots n$

$$\hat{\mu} = \sum_{i=1}^{n} \frac{x_i}{n+\gamma^2}$$

③ Let $X_1, \dots, X_n$ be independent random variables with the same probability distribution, with density f satisfying $f(\mu+x) = f(\mu-x)$ (symmetric). We assume $E[X_i] = \mu$, $Var[X_i] = \sigma^2$, and $\exists E[X_i^3]$, $i = 1..n$. Determine the value $E[(\sum_{i=1}^{n} X_i)^3]$.

Let $k \in \mathbb{N}$. The k-th moment of random variable X is $E[X^k]$.
The k-th absolute moment of random variable X is $E[|X|^k]$
The k-th central moment of random variable X is $E[(X-E[X])^k]$.
- The 1^{st} central moment of X is $E[X-E[X]] = E[X] - E[E[X]] = E[X] - E[X] = 0$
- The 2^{nd} central moment of X is $Var[X]$.

As $\exists E[X_i]$ and the distribution is symmetric with respect to μ, we can infer $E[X_i] = \mu$, so it wouldn't be needed as hypothesis.

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b+c)^3 = (a+b+c)(a+b+c)(a+b+c) = a^3 + b^3 + c^3 + 3a^2b + 3ab^2 + 3bc^2 + 3b^2c + 3c^2a + 3ca^2 + 6abc$$

$$\left(\sum_{i=1}^{n} a_i\right)^3 = \sum_{i=1}^{n} a_i^3 + 3\sum_{\substack{i \neq k \\ i,k \in \{1..n\}}} a_i^2 a_k + \sum_{\substack{i<k<l \\ i,k,l \in \{1..n\}}} a_i a_k a_l$$

$$E\left[\left(\sum_{i=1}^{n} X_i\right)^3\right] = \sum_{i=1}^{n} E[X_i^3] + 3\sum_{i \neq k} E[X_i^2 X_k] + 6\sum_{i<k<l} E[X_i X_k X_l] \overset{\text{independence}}{=}$$

$$\sum_{i=1}^{n} E[X_i^3] + 3\sum_{i \neq k} \underbrace{E[X_i^2]}_{=\mu^2+\sigma^2} \underbrace{E[X_k]}_{=\mu} + 6\sum_{i<k<l} \underbrace{E[X_i]}_{=\mu} \underbrace{E[X_k]}_{=\mu} \underbrace{E[X_l]}_{=\mu} =$$

$$n E[X_1^3] + 3(n^2-n)(\sigma^2+\mu^2)\mu + 6\binom{n}{3}\mu^3 \overset{*}{=} 3n^2\sigma^2\mu + n^3\mu^3$$

after computations

* We know the distribution of X is symmetric with respect to $\mu \Rightarrow X-\mu$ is respect to 0
$\Rightarrow (X-\mu)^3$ is respect 0 $\Rightarrow E[(X-\mu)^3] = E[X^3] - 3\mu \underbrace{E[X^2]}_{(\sigma^2+\mu^2)} + 3\mu^2 \underbrace{E[X]}_{\mu} - \mu^3 = 0$
$\Rightarrow E[X^3] = 3\mu E[X^2] - 3\mu^2 E[X] + \mu^3 = 3\mu(\sigma^2+\mu^2) - 2\mu^3 = 3\mu\sigma^2 + \mu^3$

- Another method:

• **Cummulants:** Let X, Y be independent. We assume appropiate moments exist.

$E[X+Y] = E[X] + E[Y]$. However, $E[(X+Y)^2] \neq E[X^2] + E[Y^2]$

$var[X+Y] = var[X] + var[Y]$.

We call cummulant an invariant that satisfies linearity when variables are independent.

$C_k(X) = E[W(X)]$, where W is a polinomial, $\deg W = k$.

$C_1(X) = E[X]$

$C_2(X) = var\, X$

$C_3(X) = E[(X - E[X])^3]$

$C_4(X) = E[(X - E[X])^4] - 3 var[X]^2$ (curtosis). $C_4(X+Y) = C_4(X) + C_4(Y)$

In our problem, we know $X_1 \dots X_n$ independent and

$E[(X_i - E[X])^3] = C_3(X_i)$, $i = 1 \dots n \Rightarrow$

$$C_3(\underbrace{\sum_{i=1}^n X_i}_{S}) = \sum_{i=1}^n C_3(X_i) = 0 = E[(S - E[S])^3]$$

$$E[(\sum_{i=1}^n X_i)^3] = E[S^3] = E[((S - E[S]) + E[S])^3] =$$

$$\underset{\overset{\shortparallel}{0}}{E[(S - E[S])^3]} + 3 \underbrace{E[S]}_{\overset{\shortparallel}{n\mu}} \underbrace{E[(S - E[S])^2]}_{\overset{\shortparallel}{var[S] = \sum_{i=1}^n var[X_i] = n\sigma^2}} + 3E[S]^2 \underset{\overset{\shortparallel}{0}}{E[S - E[S]]} + E[S]^3 = 3n^2\sigma^2\mu + n^3\mu^3$$

④ Let $\begin{pmatrix} X \\ Y \end{pmatrix} \rightsquigarrow N\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 9 & -6 \\ -6 & 8 \end{pmatrix}\right)$.

Compute $E[XY]$, $Var[XY]$, $COV(X^2, Y^2)$

We know $E[X^2] = Var[X] + E[X]^2$, $E[XY] = COV(X,Y) + E[X]E[Y]$

$$E[XY] = COV(X,Y) + E[X]E[Y] = -6 + 0 \cdot 0 = -6$$

$$Var[XY] = E[(XY)^2] - E[XY]^2 = E[X^2Y^2] - (-6)^2 = E[\overset{*}{X^2Y^2}] - 36 = 144 - 36 = 108$$

$$COV(X^2, Y^2) = E[X^2Y^2] - \underbrace{E[X^2]}_{\substack{Var[X] + E[X]^2 \\ 9 + 0^2 = 9}} \underbrace{E[Y^2]}_{\substack{Var[Y] + E[Y]^2 \\ 8 + 0^2 = 8}} = E[\overset{*}{X^2Y^2}] - 72 = 144 - 72 = 72$$

* we know $\begin{pmatrix} X \\ Y \end{pmatrix} \rightsquigarrow L\begin{pmatrix} U \\ V \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, where $U, V \rightsquigarrow N(0,1)$ independent and

$L \in M_2(\mathbb{R}) / LL^T = \begin{pmatrix} 9 & -6 \\ -6 & 8 \end{pmatrix}$

$$LL^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} \Rightarrow \left\{ \begin{array}{l} a^2+b^2 = 9 \\ ac+bd = -6 \\ b^2+d^2 = 8 \end{array} \right\}$$

we choose $b = 0 \Rightarrow a = \pm 3$, $d = \pm 2$, $c = -2$

$$\begin{pmatrix} X \\ Y \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 0 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 3U \\ 2V - 2U \end{pmatrix}$$

$$E[X^2Y^2] = E[(3U)^2(2U-2V)^2] = 3^2 \cdot 2^2 E[U^2(V-U)^2] = 36E[U^2V^2 - 2U^3V + U^4] =$$

$$36E[\overset{**}{U^2}]E[V^2] - 72\underbrace{E[U^3]}_{0}\underbrace{E[V]}_{0} + 36E[U^4] \overset{**}{=} 36(1 + E[U^4]) = 36(1+3) = 144 \; ***$$

Symmetric with respect to 0.

** $E[U^2] = Var[U] + E[U]^2 = 1 + 0^2 = 1$

*** $E[U^4] = \dfrac{4!}{\left(\frac{4}{2}\right)! \, 2^{4/2}} = \dfrac{4 \cdot 3 \cdot 2}{2 \cdot 2^2} = 3$

- Another approach:

We know X, Y are not independent. However, $X, \underbrace{Y - aX}_{Z}$ can be independent for appropiate $a \in \mathbb{R}$.

$$\begin{pmatrix} X \\ Z \end{pmatrix} = \begin{pmatrix} X \\ Y - aX \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \text{ has normal distribution}$$

X, Z independent $\Longleftrightarrow \operatorname{cov}(X, Z) = 0$

$$0 = \operatorname{cov}(X, Z) = \operatorname{cov}(X, Y - aX) = \operatorname{cov}(X, Y) - a\operatorname{cov}(X, X) =$$
$$\operatorname{cov}(X, Y) - a\operatorname{var}[X] = -6 - a \cdot 9 \Longleftrightarrow a = -2/3 \Longrightarrow Z = Y + \frac{2}{3}X$$ and $\begin{pmatrix} X \\ Z \end{pmatrix}$ is normally distributed with X, Z independent.

$$E[X^2 Y^2] = E\left[X^2 \left(Z - \frac{2}{3}X\right)^2\right] = E[X^2 Z^2] - \frac{4}{3}E[X^3 Z] + \frac{4}{9}E[X^4] =$$
$$\underset{\substack{\| \\ 9 + 0^2 = 9}}{E[X^2]}\, E[Z^2] - \frac{4}{3}\underset{\substack{\| \\ 0}}{E[X^3]}\, E[Z] + \frac{4}{9}\underset{\substack{\| \\ 3\operatorname{var}[X]^2 = 3 \cdot 9^2}}{E[X^4]} =$$
$$9\left(\operatorname{var}[Z] + E[Z]^2\right) + 108 = 9\operatorname{var}[Z] + 108 \overset{*}{=} 9 \cdot 4 + 108 = 144$$

$$* \ \operatorname{var}[Z] = \operatorname{var}\left[Y + \frac{2}{3}X\right] = \operatorname{var}[Y] + \frac{4}{3}\operatorname{cov}(X, Y) + \frac{4}{9}\operatorname{var}[X] =$$
$$8 + \frac{4}{3}(-6) + \frac{4}{9} \cdot 9 = 4$$

⑤ $X \rightsquigarrow N(0,1)$, $Y \rightsquigarrow N(0,3)$ independent.

Compute $P[X < Y]$, $P[X < Y-1]$, $P[|X| < |Y|]$

- $P[X<Y] = P[X-Y<0] =$

$X-Y$ has normal distribution and $E[X-Y] = E[X]-E[Y] = 0-0=0$ (symmetric with respect to 0).

$Var[X-Y] = Var[X] + Var[Y] - 2\,cov(X,Y) = 1+3-2\cdot 0 = 4$

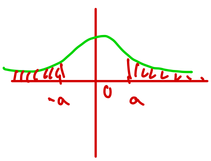
$\Longrightarrow X-Y \rightsquigarrow N(0,4)$

$X-Y$ has continuous distribution symmetric with respect to 0 $\Longrightarrow$

$P[X-Y<0] = P[X-Y>0] = 1/2$

- $P[X < Y-1] \overset{\substack{\text{standarization}\\ \uparrow}}{=} P\Big[\underbrace{\frac{X-Y}{2}}_{N(0,1)} < -\frac{1}{2}\Big] = \phi(-1/2) = 1-\phi(1/2) \approx 1-0'6915 \approx 0'3085$

ϕ: cummulative distribution function of $N(0,1)$. Value can be looked up at tables.

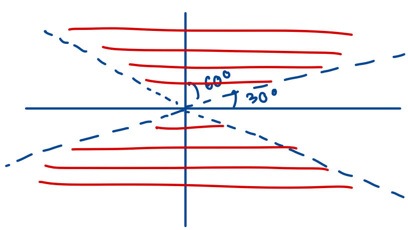


$\phi(a) + \phi(-a) = 1$

- $P[|X| < |Y|]$

Let $U = X \rightsquigarrow N(0,1)$, $V = \frac{Y-0}{\sqrt{3}} = \frac{Y}{\sqrt{3}} \rightsquigarrow N(0,1)$ independent.

$$P[|X| < |Y|] = P[|U| < \sqrt{3}\,|V|] = \iint_A f_{U,V}(u,v)\,du\,dv = \iint_A \frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}}\,du\,dv = 4/6 = 2/3$$



$A = \{(u,v) \in \mathbb{R}^2 / |u| < \sqrt{3}|v|\}$

We can see that $\equiv$ is 4/6 of the total plane.

⑥ We know $E[X_1] = \dots E[X_n] = \mu$, $Var[X_1] = \dots = Var[X_n] = \sigma^2$.
$\forall i \neq j$, $cov(X_i, X_j) = \rho\sigma^2$.
$\varepsilon_1 \dots \varepsilon_n$ independent, and also independent from $X_1 \dots X_n$ / $P[\varepsilon_i = -1] = P(\varepsilon_i = 1] = \frac{1}{2}$
$\forall i = 1 \dots n$. compute $Var[\sum_{i=1}^{n} \varepsilon_i X_i]$

We know $Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{i<j} cov(X_i, X_j)$

$$Var\left[\sum_{i=1}^{n} \varepsilon_i X_i\right] = \sum_{i=1}^{n} \underbrace{var[\varepsilon_i X_i]}_{I)} + 2\sum_{i<j} \underbrace{cov(\varepsilon_i X_i, \varepsilon_j X_j)}_{II)} = n(\mu^2 + \sigma^2)$$

ε_i, X_i independent

I) $Var[\varepsilon_i X_i] = var[\varepsilon_i^2 X_i^2] = E[\varepsilon_i^2 X_i^2] - E[\varepsilon_i X_i]^2 =$
$\underbrace{E[\underbrace{\varepsilon_i^2}_{=1}]}_{=1} E[X_i^2] - (E[\varepsilon_i]E[X_i])^2 = E[X_i^2] = Var[X_i] + E[X_i]^2 = \sigma^2 + \mu^2$, with $E[\varepsilon_i] = 0$

II) $cov(\varepsilon_i X_i, \varepsilon_j X_j) = E[\varepsilon_i \varepsilon_j X_i X_j] - E[\varepsilon_i X_i]E[\varepsilon_j X_j] =$
$E[\varepsilon_i]E[\varepsilon_j]E[X_i X_j] - E[\varepsilon_i]E[X_i]E[\varepsilon_j]E[X_j] = 0$, with $E[\varepsilon_i] = 0$, $E[\varepsilon_j] = 0$

⑦ Let $X_1 \dots X_{10}$, $X_{11} \dots X_{20}$ independent, $X_1 \dots X_{10} \rightsquigarrow N(\mu_1, \sigma_1^2)$, $X_{11} \dots X_{20} \rightsquigarrow N(\mu_2, \sigma_2^2)$
let $\bar{X} = \frac{1}{20}\sum_{i=1}^{20} X_i$, $\bar{X}_1 = \frac{1}{10}\sum_{i=1}^{10} X_i$, $\bar{X}_2 = \frac{1}{10}\sum_{i=11}^{20}$
Find $\alpha, \beta \in \mathbb{R}$ / the estimator $\hat{\sigma}^2 = \alpha\sum_{i=1}^{20}(X_i - \bar{X})^2 + \beta(\bar{X}_1 - \bar{X}_2)^2$ is unbiased.
$(\sigma^2 = E[\hat{\sigma}^2])$

$$\sigma^2 = \hat{\sigma}^2 = \alpha\sum_{i=1}^{20} \underbrace{E[(X_i - \bar{X})^2]}_{II)} + \beta \underbrace{E[(\bar{X}_1 - \bar{X}_2)^2]}_{I)}$$

I) $E[(\bar{X}_1 - \bar{X}_2)^2] = Var[\bar{X}_1 - \bar{X}_2] + E[\bar{X}_1 - \bar{X}_2]^2 = Var[\bar{X}_1] + Var[\bar{X}_2] + (\mu_1 - \mu_2)^2 =$ Independent
$\frac{\sigma^2}{10} + \frac{\sigma^2}{10} + (\mu_1 - \mu_2)^2 = \frac{\sigma^2}{5} + (\mu_1 - \mu_2)^2$

- $i \leq 10$

$$E[(X_i-\bar{X})^2] = \text{var}[X_i-\bar{X}] + E[X_i-\bar{X}]^2 =$$

$$\text{var}[X_i] - 2\,\text{cov}(X_i,\bar{X}) + \text{var}[\bar{X}] + (\mu_1 - E[\bar{X}])^2 =$$

$$\sigma^2 - \frac{2}{20}\underbrace{\sum_{j=1}^{20}\text{cov}(X_i,X_j)}_{=\sigma^2} + \frac{1}{20^2}\underbrace{\text{var}[\sum_{j=1}^{20}X_j]}_{=20\sigma^2} + \mu_1 - \Big(\underbrace{\sum_{j=1}^{10}E[X_j]}_{=10\mu_1} + \underbrace{\sum_{j=11}^{20}E[X_j]}_{=10\mu_2}\Big)^2 =$$

$$\sigma^2 - \frac{1}{10}\sigma^2 + \frac{1}{20}\sigma^2 + (\mu_1 - \tfrac{1}{2}\mu_1 - \tfrac{1}{2}\mu_2)^2 = \frac{19}{20}\sigma^2 + \frac{1}{4}(\mu_1-\mu_2)^2$$

- $10 < i \leq 20$ Analog

$$\sigma^2 = \alpha\,(19\sigma^2 + 5(\mu_1-\mu_2)^2) + \beta\left(\frac{\sigma^2}{5} + \mu_1+\mu_2\right)$$

coeffitient near $(\mu_1-\mu_2)^2$

coeffitient near σ^2

$$\left.\begin{aligned} 5\alpha + \beta &= 0 \\ 19\alpha + \frac{\beta}{5} &= 1 \end{aligned}\right\} \begin{aligned} \alpha &= 1/18 \\ \beta &= -5/18 \end{aligned}$$

Another approach:

$$\frac{1}{9}\sum_{i=1}^{10}(X_i-\bar{X}_1)^2, \quad \frac{1}{9}\sum_{i=11}^{20}(X_i-\bar{X}_2)^2$$

$$\frac{1}{18}\sum_{i=1}^{10}(X_i-\bar{X}_1)^2 + \frac{1}{18}\sum_{i=11}^{20}(X_i-\bar{X}_2)^2 =$$

$$\frac{1}{18}\sum_{i=1}^{10}\left(X_i-\bar{X} + (\bar{X}-\bar{X}_1)^2\right) + \frac{1}{18}\sum_{i=11}^{20}\left(X_i-\bar{X} + (\bar{X}-\bar{X}_2)\right)^2 =$$

$$\frac{1}{18}\sum_{i=1}^{10}(X_i-\bar{X})^2 + \frac{1}{18}\cdot 2\cdot\sum_{i=1}^{10}(X_i-\bar{X})\left(\frac{\bar{X}_2-\bar{X}_1}{2}\right) + \frac{10}{18\cdot 2^2}(\bar{X}_2-\bar{X}_1)^2 +$$

$$\frac{1}{18}\sum_{i=11}^{20}(X_i-\bar{X})^2 + \frac{1}{18}\cdot 2\cdot\sum_{i=11}^{20}(X_i-\bar{X})\left(\frac{\bar{X}_1-\bar{X}_2}{2}\right) + \frac{10}{18\cdot 2^2}(\bar{X}_1-\bar{X}_2)^2 =$$

$$\frac{1}{18}\sum_{i=1}^{20}(X_i-\bar{X})^2 + \frac{5}{18}(\bar{X}_1-\bar{X}_2)^2 + \frac{1}{18}(\bar{X}_2-\bar{X}_1)(10\bar{X}_1 - 10\bar{X}) +$$

$$\frac{1}{18}(\bar{X}_1-\bar{X}_2)(10\bar{X}_2 - 10\bar{X}) =$$

$$\frac{1}{18}\sum_{i=1}^{20}(X_i-\bar{X})^2 + \frac{5}{18}(\bar{X}_1-\bar{X}_2)^2 - \frac{5}{18}(\bar{X}_1-\bar{X}_2)^2 - \frac{5}{18}(\bar{X}_1-\bar{X}_2)^2 =$$

$$\frac{1}{18}\sum_{i=1}^{20}(X_i-\bar{X})^2 - \frac{5}{18}(\bar{X}_1-\bar{X}_2)^2$$

8 $X_1, \ldots X_{100} \rightsquigarrow N(\mu, \sigma^2)$ independent. μ is unknown, σ^2 known.

We constructed a satisfied test of $H_0: \mu = \mu_0$ against $H_1: \mu > \mu_0$ on significance level 0'05. $\forall i \neq j$ $corr(X_i, X_j) = \frac{1}{10}$.

Find the real size of the constructed test.

$\mu \approx \hat{\mu} = \overline{X}$

We should reject H_0 when $\overline{X} > a$, $a \in \mathbb{R}$ a?

We want the following:

If H_0 is satisfied, $(X_1 \ldots X_{100} \rightsquigarrow N(\mu_0, \sigma^2))$

$$P[\neg H_0] = P[\underset{N(\mu_0, \frac{\sigma^2}{100})}{\overline{X}} > a] = 0'05 = P[\underset{N(0,1)}{\frac{\overline{X}-\mu_0}{\sigma/10}} > \frac{a-\mu_0}{\sigma/10}] \Leftrightarrow P[\frac{\overline{X}-\mu_0}{\sigma/10} \leq \frac{a-\mu_0}{\sigma/10}] = 0'95$$

table

$$\Rightarrow \frac{a-\mu_0}{\sigma/10} = 1'645 \Rightarrow a = \mu_0 + 0'1645\,\sigma$$

We reject $H_0 \Leftrightarrow \mu_0 + 0'1645\,\sigma < \overline{X}$

We change the distribution of $X_1 \ldots X_n$ / $corr(X_i, X_j) = \frac{1}{10}$ $\forall i \neq j$

We assume H_0 is satisfied $(\mu = \mu_0)$ and we are requested to compute

$$P[\neg H_0] = P[\overline{X} > \mu_0 + 0'1645\,\sigma]$$

$$corr(X,Y) = \frac{COV(X,Y)}{Var[X]\,Var[Y]}$$

$$Var[\overline{X}] = \frac{1}{100^2} Var\left[\sum_{i=1}^{100} X_i\right] = \frac{1}{100^2}\left(\sum_{i=1}^{100} Var[X_i] + 2\sum_{i<j} COV(X_i, X_j)\right) \overset{\uparrow}{=}$$

$$\frac{1}{100^2}\left(100 \cdot \sigma^2 + 100 \cdot 99 \cdot \frac{1}{10} \cdot \sigma \cdot \sigma\right) = \frac{\sigma^2}{100} + \frac{99}{1000}\sigma^2 = \frac{109}{1000}\sigma^2 = 0'109\,\sigma^2$$

$$\overline{X} \rightsquigarrow N(\mu_0, 0'109\,\sigma^2)$$

$$P[\overline{X} > \mu_0 + 0'1645\,\sigma] = P\left[\frac{\overline{X}-\mu_0}{\sigma\sqrt{0'109}} > \frac{0'1645}{\sqrt{0'109}}\right] =$$

$$1 - P\left[\frac{\overline{X}-\mu_0}{\sigma\sqrt{0'109}} \leq 0'498\right] \approx 1 - 0'69 \approx 0'31$$