Exercises I2

Exercise 2.1

i) $S^2 \leq \frac{1}{n-1} \sum\limits_i^n (X_i - a)^2$ for any $a \in \mathbb{R}$

$S^2 = \frac{1}{n-1} \sum\limits_i^n (X_i - \overline{X})^2 \leq \frac{1}{n-1} \sum\limits_i^n (X_i - a)^2$

$\sum\limits_i^n (X_i - \overline{X})^2 \leq \sum\limits_i^n (X_i - a)^2$

$\sum\limits_i^n (-2X_i\overline{X} + \overline{X}^2) \leq \sum\limits_i^n (-2X_ia + a^2)$

$\sum\limits_i^n (-2X_i\overline{X}) + n\overline{X}^2 \leq -2a\sum\limits_i^n(X_i) + \sum\limits_i^n(a^2)$

$-2n\overline{X}^2 + n\overline{X}^2 \leq -2an\overline{X} + na^2$

$-n\overline{X}^2 \leq -2an\overline{X} + na^2$

$-\overline{X}^2 \leq -2a\overline{X} + n^2a^2$

$0 \leq \overline{X}^2 - 2a\overline{X} + n^2a^2 (\star)$

if a and $\overline{X}$ have the opposite sign the inequality $(\star)$ is true because a sum of positive numbers is ≥ 0.

otherwise:

$\overline{X}^2 - 2a\overline{X} + n^2a^2 \geq^{(n\geq 1)} \overline{X}^2 - 2an\overline{X} + n^2a^2 = (\overline{X} - na)^2 \geq 0$

$\square$

ii) $\frac{(n-1)S^2}{n} = \frac{1}{n}\sum_i^n (X_i - \overline{X})^2$

$= \frac{1}{n}\sum_i^n (X_i^2 - 2X_i\overline{X} + \overline{X}^2)$

$= \overline{X^2} - 2\overline{X}^2 + \overline{X}^2 = \overline{X^2} - \overline{X}^2$

$\square$

Exercise 2.2

$X_1, \ldots, X_n$ iid random sample with $X_i \sim F$ with continuous F

.

i)

$\min : \mathbb{R}^n \to \mathbb{R}$

I call Z the rv $X_{(1)}$

$B_z = \{x_1, \ldots, x_n : \min(x_1, \ldots, x_n) \leq z\}$

we have that B_Z is the set of points in which at least one component is less than z

${B_z}^C = (z, +\infty) \times \cdots \times (z, \infty)$

$F_Z(z) = 1 - \int\limits_{B_z^C} \prod_{i=1}^n f(x_i)dx_1dx_2\cdots dx_n = 1 - (1 - F(z))^n$

Now we can take the derivative:

$f_Z(z) = -n(1 - F(z))^{n-1}(-f(z)) = n(1 - F(z))^{n-1}f(z) \quad \square$

ii)

$\max : \mathbb{R}^n \to \mathbb{R}$

$B_z = \{x_1, \ldots, x_n : \max(x_1, \ldots, x_n) \leq z\}$

which means that B_z is the set of points in which all components are smaller than z (definition of minimum)
$B_z = (-\infty, z] \times \cdots \times (-\infty, z]$

we can find the df of Z,
$F_Z = \int\limits_{B_z} \prod\limits_{i=1}^{n} f(x_i) dx_1 dx_2 \cdots dx_n = (F(z))^n$
Now we can take the derivative:
$f(z) = n(F(z))^{n-1} f(z) \quad \square$
iii) if they are not independent we cannot factorize the pdfs and we would need the joint pdf to be able to calculate the integrals.
iv) Bz??????

Exercise 2.5

discrete

- Bernoulli: coin toss that either gives heads (1) with probability θ or tails (0) with probability $1 - \theta$
- Binomial: number of successes in n independent trials (that either succeed or fail) each with probability θ, for example number of heads when tossing a coin n times
- NegBin: it's a generalized version of geometric rv that models the number of failures until r successes happen (in n independent binary trials like the binomial). For

example the number of trials until we get r non consecutive heads when tossing a coin
- Poisson: number of events that happen independently from each other in a fixed amount of time. For example the number of connections to a server in a second.

continuous

- Gaussian: for example the height of humans.
- Exponential: for example the time a client waits in queue before being served by a server.
- Gamma: it's a generalization of the exponential distribution. It is also used to model waiting times.
- Weibull: for example it can model the time an electronic device lasts.
- Uniform: in telecommunications it can model the granular error of a symmetrical quantizer.