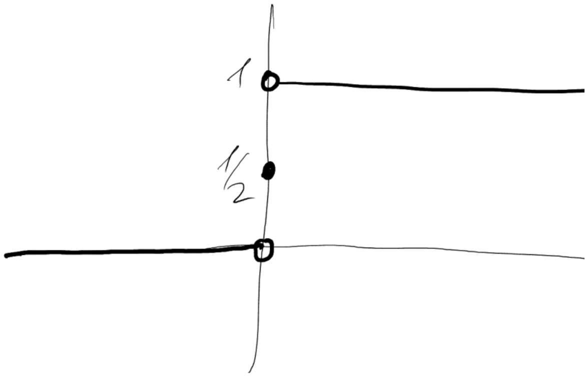
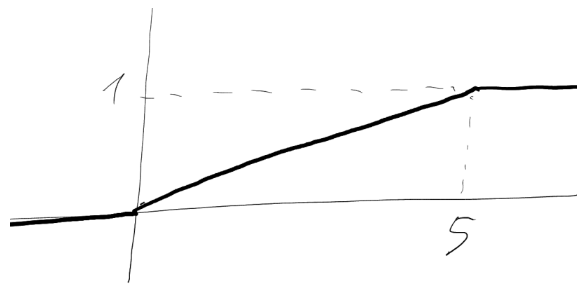
Solutions to 1st exam NMAI059 Probability and Statistics 1 – June 16, 2021

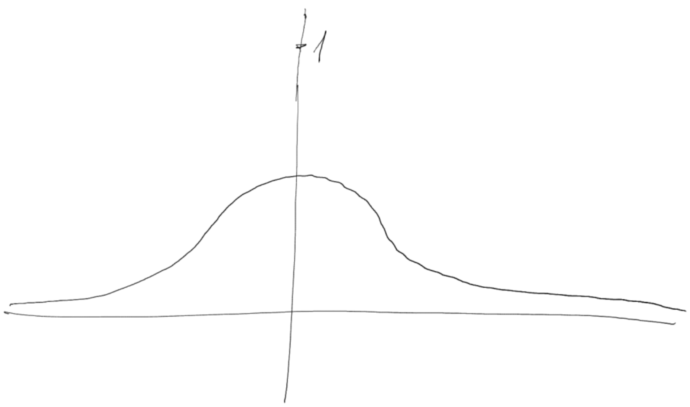
1. (10 points) (a) Decide, which of the figures describe a cdf of some random variable. Do the next two parts just for those, that show a cdf.

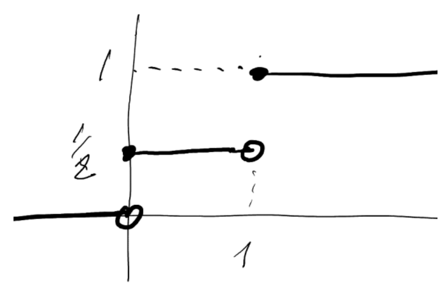
(b) Estimate the expectation.

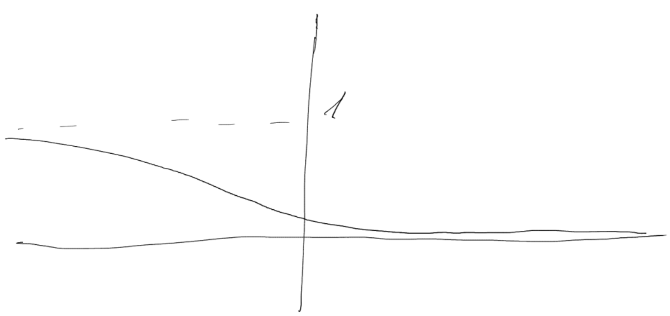
(c) Order the variables by their variance.

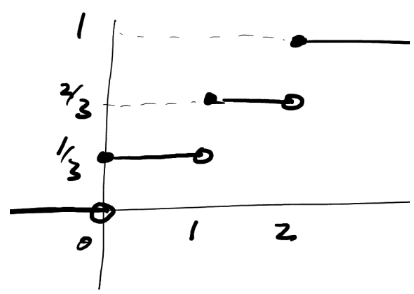


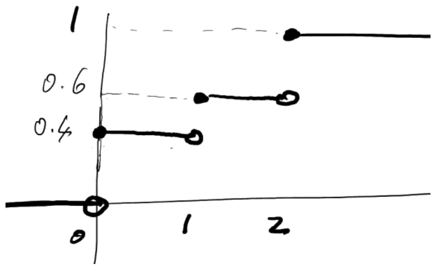












Solution:

(a) The third and the fifth function are not cumulative distribution functions of any random variable because they are non-decreasing. Neither is the first, because it is not right continuous. (Or: the relevant random variable would have to with probability 1/2 generate numbers that are positive but less than any positive real number.) All the other functions meet all the requirements from the lecture (non-decreasing, right-continuous, proper limits in $\pm\infty$), so they are cdf's. The second function corresponds to the uniform distribution $U(0,5)$, the fourth, sixth and seventh correspond to a discrete random variable that takes only the values in $\{0,1\}$ or in $\{0,1,2\}$; the probability of a value is equal to the size of the jump in the cdf at that point. We denote next by X_i the random variable whose cdf is in the i-th figure.

(b) $X_2 \sim U(0,5)$, thus $\mathbb{E}(X_2) = 5/2$. (In case we forgot this, we could compute it as $\int_0^5 x\frac{1}{5}dx = \frac{1}{5}\frac{5^2}{2}$.)
$X_4 \sim Ber(1/2)$, so $\mathbb{E}(X_4) = 1/2$
For $X = X_6, X_7$ we observe that $P(X=0) = P(X=2)$ and so $\mathbb{E}(X_6) = \mathbb{E}(X_7) = 1$. (To check: $P(X_6 = 0) = 1/3$, $P(X_7 = 0) = 0.4$.)

(c) $X_2 \sim U(0,5)$, therefore $\mathrm{var}(X_2) = 5^/12$.
$X_4 \sim Ber(1/2)$, consequently $\mathrm{var}(X_4) = 1/4$
For $X = X_6, X_7$ we have $P(X=0) = P(X=2)$ and $\mathbb{E}(X) = 1$. Therefore $\mathrm{var}(X) = \mathbb{E}((X-1)^2) = 1 \cdot P(X \neq 1)$.
Conclusion: $\mathrm{var}(X_4) < \mathrm{var}(X_6) < \mathrm{var}(X_7) < \mathrm{var}(X_2)$.

---

2. (10 points) (a) Two players, Adam and Beatrix, roll the dice repeatedly, in order ABABAB... What is the probability that Adam will roll a six first?

(b) Cecil joins the game, they now roll in the order ABCABCABC... The probability that Adam will get the six first, then Beatrix, and only after that Cecil is 216/1001. Explain. (If Adam gets a six more than once, and only then Beatrix, that's fine too, we're just concerned with the order of the first time they roll a six.)

Solution:

(a) Let p be the desired probability. We will use the total probability formula. For the desired outcome there are two possibilities: either Adam gets the six the first time (probability 1/6) and it doesn't matter what happens next, or neither Adam nor Beatrix roll a six (probability $(5/6)^2$) and we're in the same situation as at the beginning, i.e. the probability of success is p. Overall, then, we know that

$$p = \frac{1}{6} + \left(\frac{5}{6}\right)^2 \cdot p$$

From here, a simple calculation gives $p = 6/11$.

Alternatively, we can label as A_i the event that Adam rolls a six on the $(i+1)$th roll, and the previous $2i$ rolls never resulted in a six. Obviously $P(A_i) = \frac{1}{6}(5/6)^{2i}$ and we are interested in

$$p = P\Big(\bigcup_{i=0}^{\infty} A_i\Big) = \frac{1}{6}\sum_{i=0}^{\infty}\big(\frac{5}{6}\big)^{2i}.$$

The sum of the geometric series gives us the same result $6/11$.

(b) We will proceed similarly to the previous part, this time denoting the probability we are looking for by q. Again, we have the possibility of a turn where Adam, Beatrix, and Cecil do not get a six (probability $(5/6)^3$) and we are in the same situation as at the beginning, i.e., we then have a probability of success q. If Adam gets a six the first time, it doesn't matter what he gets next, and we can take him out of the consideration. Beatrix and Cecil then play the game from part (a), where the probability of success is $p = 6/11$. Overall, then, we have

$$q = \frac{1}{6} \cdot \frac{6}{11} + \left(\frac{5}{6}\right)^3 \cdot q$$

Easy calculation leads to $q = 216/1001$.

---

3. (10 points) Pareto distribution with parameter $\alpha > 1$ has pdf $f(x) = \frac{\alpha}{x^{\alpha+1}}$ for $x \in [1, \infty)$ (and zero elsewhere).

(a) Verify that it is a pdf.

(b) We sample values 5, 2, 3 from this distribution. Derive a point estimate $\hat{\alpha}$ using the maximal likelihood method.

(c) Let X follow the Pareto distribution, that is let $f_X = f$. Calculate $\mathbb{E}(X)$.

(d) Find the point estimate $\hat{\alpha}$ by the moment method.

Solution:

(a) We need to verify that the integral over all x is 1:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{1}^{\infty} \alpha x^{-\alpha-1}dx = \big[\frac{\alpha x^{-\alpha}}{-\alpha}\big]_1^{\infty} = 0 - (-1) = 1.$$

(b) The probability of the triple $(5, 2, 3)$ is

$$L(5, 3, 2; \alpha) = \frac{\alpha}{5^{\alpha+1}}\frac{\alpha}{2^{\alpha+1}}\frac{\alpha}{3^{\alpha+1}}.$$

So we want to maximize this function. For simplicity, we first logarithm it, then use the derivative:

$$\begin{aligned}\ell(\alpha) &= \big(\log\alpha - (\alpha+1)\log 5\big) + \big(\log\alpha - (\alpha+1)\log 2\big) + \big(\log\alpha - (\alpha+1)\log 3\big)\\ \ell'(\alpha) &= \frac{3}{\alpha} - \log(5\cdot 2\cdot 3)\end{aligned}$$

The maximum likelihood estimate is such α for which $\ell'(\alpha) = 0$ (we see that for smaller αs the derivative is positive, for larger ones negative, so it is indeed a maximum). So we put $\hat{\alpha} = \frac{1}{3}\log 30 \doteq 1.13$. (No need to numerically evaluate this in the test.)

(c) By definition, we have

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\cdot f(x)dx = \int_{1}^{\infty} \alpha x^{-\alpha}dx \quad = \Big[\frac{\alpha}{-\alpha+1}x^{-\alpha+1}\Big]_1^{\infty} = 0 - \Big(\frac{\alpha}{-\alpha+1}\Big) \quad = \frac{\alpha}{\alpha-1}$$

Note that the 0 on the second line is $\lim_{x\to\infty} x^{-\alpha+1}$. So here we take advantage of the fact that $\alpha > 1$.

(d) We use the previous part. The first moment of a given distribution is $\alpha/(\alpha-1)$, we set this equal to the first selection moment $\bar{x}_3 = (5+2+3)/3 = 10/3$. So we solve the easy equation $\alpha/(\alpha-1) = \bar{x}_3$, from here $1 - 1/\alpha = 1/\bar{x}_3$, so the estimate obtained by the moment method is

$$\hat{\alpha} = \frac{1}{1 - 1/\bar{x}_3} = \frac{\bar{x}_3}{\bar{x}_3 - 1} = \frac{10}{7} \doteq 1.43.$$

---

4. (10 points) (a) Define the notion joint cumulative distribution function.

(b) Describe, how to compute the empirical cumulative distribution function.

Solution:

(a) For the random variables $X_1, \ldots, X_n$ their joint cumulative distribution function is defined by the formula

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P(X_1 \le x_1 \,\&\ldots\&\, X_n \le x_n).$$

(The definition for two random variables was sufficient for the test: $F_{X,Y}(x,y) = P(X \le x \,\&\, Y \le y)$.)

(b) If we have a random sample $X_1, \ldots, X_n$, then the corresponding empirical distribution function is

$$\hat{F}_n(x) = \frac{\sum_{i=1}^n I(X_i \le x)}{n},$$

where $I(X_i \le x) = 1$ if $X_i \le x$ and 0 otherwise.

Equivalently, if we sample $x_1, \ldots, x_n$ (the realization of the random sample), then the empirical distribution function $\hat{F}(x)$ is one n-th of the number of i's such that $x_i \le x$.

---

5. (10 points) State the Central limit theorem. Explain what is it good for.

Solution: Let $X_1, X_2, \ldots$ be equally distributed random variables with mean μ and variance σ^2. Put

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then Y_n converge in distribution to $N(0,1)$; that is, for each $x \in \mathbb{R}$

$$\lim_{n\to\infty} P(Y_n \le x) = \Phi(x).$$

The significance of the theorem is that it allows us to approximate the sum of many random variables by a single, well-known distribution. Such sums are often found – for example, $Bin(n,p)$ is the sum of n independent variables following $Ber(p)$, and many physical phenomena are (approximately) described as the sum of of independent random variables. This explains why binomial coefficients, like many quantities that occur in practice, have approximately normal distributions.

---

6. (10 points) State and prove the theorem about expectation of a sum of random variables. (Only prove it for the case of discrete random variables.)

Solution:

Theorem: $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, if the right-hand side is defined. (That is, both expectations are defined and it is not the case $\infty - \infty$.)

Proof:

We use LOTUS for $g(x,y) = x + y$. We get

$$\begin{aligned}
\mathbb{E}(g(X,Y)) &= \sum_{x\in ImX} \sum_{y\in ImY} g(x,y) \cdot P(X = x \,\&\, Y = y) \\
&= \sum_{x\in ImX} \sum_{y\in ImY} (x+y) \cdot P(X = x \,\&\, Y = y) \\
&= \sum_{x\in ImX} \sum_{y\in ImY} x \cdot P(X = x \,\&\, Y = y) + \sum_{x\in ImX} \sum_{y\in ImY} y \cdot P(X = x \,\&\, Y = y) \\
&= \sum_{x\in ImX} x \cdot P(X = x) + \sum_{y\in ImY} y \cdot P(Y = y) \\
&= \mathbb{E}(X) + \mathbb{E}(Y)
\end{aligned}$$

Here we used that $\sum_{y\in ImY} P(X = x \,\&\, Y = y) = P(X = x)$, and similarly when X and Y are switched.