PROBABILITY THEORY

MATH 154

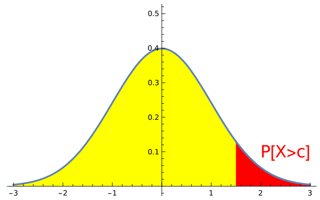
Homework 5

Problem 5.1: a) Formulate Jensen inequality in the case $f(x) = |x|$. and use it to verify calculus identity $|\int_0^1 f(x)\ dx| \leq \int_0^1 |f(x)|\ dx$ for a continuous function on $[0, 1]$.
b) Assume $X \in \mathcal{L}^4$ that has mean 0 and variance 1. Prove that the kurtosis is larger or equal than 1.
c) We have seen in class that Jensen can prove the geometric-arithmetic mean inequality $\sqrt{ab} \leq (a + b)/2$. Prove in general for $a_i \geq 0$ that $(\prod_{i=1}^n a_i)^{1/n} \leq (\sum_{i=1}^n a_i)/n$.

Problem 5.2: A probability space and random variable X defines what one calls a **null hypothesis**, the assumption that an effect does not exist. Assume you measure $X = c$ and that c is larger than the expectation, then the P**-value** of this experiment is defined as $\mathrm{P}[X \geq c]$. If the P-value is < 0.05, one considers the result as **significant** and rejects the null-hypothesis. If the P-value is > 0.05, one fails to reject the null hypothesis.
a) Assume a hypothesis is that X is exponentially distributed. You measure $X = 2$. What is the p-value?
b) Estimate the p-value using Chebyshev's inequality.
c) Having a p-value smaller than 0.05 is considered the gold standard for "statistical significance". Discuss the following strategy: we repeat an experiment a couple of times until the P-value is smaller than 5 percent. You label the early runs as warm-up-test runs and publish the paper.
d) Is it true that if you make a measurement and see the P value is larger than 0.05 that the non-significance means that the effect does exist? Explain in an example.

[1]

[1] If c was smaller than the expectation, we would define the P-value as $\mathrm{P}[X \leq c]$.



Problem 5.3: You own an insurance company that gets random claims at random times. In order to have enough reserves, you want to estimate how large claims will be in the future. Your staff tells you the mean and standard deviation of the historical claim distribution but you do not know the distribution.
a) Why does Chebyshev's inequality imply that at least 89 percent of future claims will be within three standard deviations away from the mean?
b) Build a similar rule of thumb to see that [ ] percent of future claims are within two standard deviations from the mean. Explain.
c) Fill in the box: 96 percent of future claims are within [ ] standard deviations from the mean. Explain.

Problem 5.4: We study first entropy $S(\mathcal{A})$ for a finite σ-algebra, where random variables are just step functions.
a) Show that $f(x) = x\log(x)$ is convex. The limit $f(0) = 0$ exists.
b) Show: the uniform distribution on $\{1, \dots, n\}$ has maximal entropy.
c) Let $\mathcal{A}_X$ be the σ-algebra of a random variable $X \in \mathcal{S}$ and $\mathcal{A}_{X,Y}$ the σ algebra of two random variables $X, Y \in \mathcal{S}$. Show that if X, Y are independent, then $S(\mathcal{A}_{X,Y}) = S(\mathcal{A}_X) + S(\mathcal{A}_Y)$.

Problem 5.5: The **differential entropy** of a random variable with continuous distribution $f = f_X$ is defined as $S = -\int_{-\infty}^{\infty} f(x)\log(f(x))\, dx$. The understanding is to put $u\log(u) = 0$ if $u = 0$ and that f is such that S exists.
a) Compute the differential entropy for the uniform distribution on $[0, 1/2]$.
b) Use a convex function to show that $S \leq \log(|K|)$, where $K = \{x \in \mathbb{R}, f(x) > 0\}$ is the support of the density function and $|K| = \int_K 1\, dx$.
c) How come that the Shannon entropy $-\sum_i p_i \log(p_i)$ is non-negative, while the differential entropy can become negative?
d) Compute the differential entropy of the standard normal distribution. What happens if we translate or scale a random variable? Assume you know $S(X)$. What is $S(X + c)$? What is $S(\lambda X)$. Use this to get the entropy of a general Gaussian random variable.
e) Use Jensen to verify that $S(X) \geq -\log \int f_X(x)^2\, dx$. The right hand side is called the **collision entropy** or **Rényi entropy**.