1 Probability and random variables

- Probability: S sample space (all possible states of the system), $F \subset \mathcal{P}(S)$ a σ-algebra, $P : F \to \mathbb{R}$ a measure, such that $P(S) = 1$.
- Random variable: $X : S \to \mathbb{R}$, such that preimages of open sets are in F (i.e. has a well defined probability).
- (Cumulative) distribution function of random variable: $F_X(t) = P(X \leq t)$.
- Probability distribution of random variable: g such that $F_X(t) = \sum_{x \leq t, x \in C} g(x)$.
- Probability density function: f such that $F_X(t) = \int_{-\infty}^{t} f(s)ds$.
- Two random variables have the same distribution if they have the same cdf.

Example: uniform distribution:

- S a finite interval $[a, b]$
- F: Set of Borel sets on S (sets with a well defined "length")
- P: Borel measure ("length") divided by $b - a$
- $X = id$.

1.1 Expectation of random variables and their functions

- X is a random variable, the expectation of X is $E[X] = \int_S X dP$.
- The variance of X is $E[(X - E[X])^2]$.
- The k-th moment of X is $E[X^k]$.
- The moment generating function of X is $E[e^{Xt}]$ (two sided Laplace transform)
- The characteristic function of X is $E[e^{itX}]$ (Fourier transform)

Since expectation is defined via integration, one can use the properties of integration to prove statements regarding expectation.

Example: Chebyshev's theorem: $E[X] = 0$, $E[X^2] = 1$, then $P(|X| < k) \geq 1 - \frac{1}{k^2}$. Proof:

$$1 = E[X^2] = \int_S X^2 dP \geq k^2 \int_{|X| \geq k} 1 dP = k^2(1 - P(|X| < k))$$

Example: If X has pdf f_X, then $E[g(X)] = \int_{-\infty}^{\infty} g f_x dt$. We prove it when $g(X)$ is bounded via Fubini's theorem:

$$E[g(X)] = \int_S g(X) dP$$

$$= \int_{g(X)\geq 0} \int_0^{g(X)} 1 dy dP - \int_{g(X)<0} \int_{g(X)}^0 1 dy dP$$

$$= \int_0^{\infty} \int_{g^{-1}([y,\infty])} f_X(t) dt dy - \int_{-\infty}^0 \int_{g^{-1}([-\infty,y])} f_X(t) dt dy$$

$$= \int_{-\infty}^{\infty} g f_x dt$$

(There is a multivariate version of this formula)

Can you write down a random variable with neither probability distribution nor p.d.f.?

Can you write down a random variable with no expectation?

1.2 Independence of random event, conditional probability

- $A, B \in F$ are independent iff $P(A \cap B) = P(A)P(B)$.
- If $P(B) \neq 0$, $P(A \cap B) = P(B)P(A|B)$. Here $P(A|B)$ is the conditional probability of A when B is known to happen.

1.3 Independence of random variables, conditional distribution

- X and Y are two random variables. The joint (cumulative) distribution function is $F(s,t) = P(X \leq s, Y \leq t)$
- If $F(s,t) = \sum_{(x,y)\in C, x\leq s, y\leq t} g(s,t)$, we call g the joint probability distribution.
- If $F(s,t) = \int_{(-\infty,s]\times(-\infty,t]} f(x,y) dx dy$ we call f the joint probability density.
- X and Y are called independent iff the joint cdf is $F(x,y) = F_X(x)F_Y(y)$.
- Knowing the joint distribution of X and Y, the distribution of X or Y are called the marginal distribution, their p.d. or p.d.f. the marginal p.d. or marginal p.d.f.

- If X and Y has a "good" joint probability sensity f, we can define conditional distribution of X at $Y = y$ as the one with desnity $\frac{f(x,y)}{h(y)}$ where h is the marginal p.d.f $h(y) = \int_{\mathbb{R}} f(x,y)dx$.
- The covariance between X and Y is $E[(X - E[X])(Y - E[Y])]$

Example: X and Y are two independent random variable with uniform distribution on $[0,1]$. What is the joint distribution function of X and Y? How about $max(X,Y)$ and $min(X,Y)$? What are their covariances?

2 Special probability distributions, central limit theorem

3 Sample statistics

4 Point estimators and their properties

5 Method of moments, Maximum likelihood

6 Maximum a posteriori

7 Hypothesis testing

8 Examples of hypothesis testing

9 Confidence interval

10 Linear Regression

11 ANOVA

12 Example of non parametric methods