Lecture 6 - Moments of Random Variables

Functions of Random Variables

- A measurable real-valued function $Y = g(X)$ of a random variable X is also a random variable
- The distribution of a function of a random variable is defined as:

$$P_Y(A) = P_X(g^{-1}(A))$$

- or in the event notation:

$$P(Y \in A) = P(X \in g^{-1}(A))$$

- If g is a discrete function, i.e. $X \in \{x_1, x_2, \ldots\}$ and $Y = g(X) \in \{y_1, y_1, \ldots\}$ then:

$$P(Y = y) = \sum_{x:g(x)=y} P(X = x)$$

- In other words, to calculate the probability of $P(Y = y)$ we sum probabilities $P(X = x)$ for all x such that $g(x) = y$.

Expected value

- Let $X \in \{x_1, x_2, \ldots\}$ be a discrete random variable. The expected value (expectation, mean value) of X is a number:

$$E(X) = \sum_i x_i \cdot P(X = x_i)$$

- We will often drop the parentheses and denote $E(X)$ as EX.

- If X is always larger or equal to a then $EX \geq a$.
- If X is always smaller or equal to b then $EX \leq b$.
- Alternatively we can define expected value for a random variable that takes values only within nonnegative integers as:

$$EX = \sum_{k=1}^{\infty} P(X \geq k)$$

- A random variable might not have a finite expectation, i.e. it might diverge to infinity.
- We can calculate the expected value of a function of a random variable by:

$$E(g(X)) = \sum_{x} g(x) \cdot P(X = x)$$

- Expectation is linear:

$$E(aX + b) = aE(X) + b$$

- In particular:
  - The expected value of any number b is $E(b) = b$
  - Hence, $E(EX) = EX$
  - A constant multiplying X can be taken outside the expectation: $E(aX) = a \cdot EX$

Distribution	Expected Value
Bernoulli	$EX = p$
Binomial	$EX = np$
Geometric	$EX = \frac{1}{p}$
Negative Binomial	$EX = \frac{rp}{1-p}$
Poisson distribution	$EX = \lambda$

St. Petersburg paradox

- Consider a game in which we toss a fair coin until the first head. The winning in this game is 2^k dollars, where k - number of tosses. How much entry fee are we willing to pay to play this game?
- X — number of tosses till first head, $X \sim G_1(\frac{1}{2})$:
- $P(X = k) = (1-p)^{k-1}p = (\frac{1}{2})^k$
- $Y = g(X) = 2^X$ — winning
- We are willing to pay a dollars as long as $EY - a > 0$.
- $EY = \sum_{k=1}^{\infty} 2^k (\frac{1}{2})^k = 1 + 1 + 1 + \ldots = \infty$
- Are we willing to pay any amount of the entry fee???

Utility theory (optional)

- Utility theory explain the behaviour of people in such games (as in St. Petersburg paradox)
    - People do not decide based on the expected winning, but rather the expected utility of winning
    - The utility grows much slower than the winning itself
- Example: consider a game in which with probability 20% we get 10 times the amount of bet, while with probability 80% we lose the bet. The expected winning is 1.2 times the money we bet. Would you bet the savings of your entire life in this game?
- To analyse this example we first need to choose our utility function. A popular choice is the logarithmic utility function:

$$U(y) = \log_2(y)$$

- In the previous example having initial capital y_0 and betting a dollars, the expected utility is:
- $EU = 0.2 \cdot \log_2(y_0 + 10a) + 0.8 \cdot \log_2(y_0 - a)$

- This the expected utility increase is:

$$\Delta U = EU - U(y_0) = 0.2 \cdot \log_2(y_0 + 10a) + 0.8 \cdot \log_2(y_0 - a) - \log_2(y_0)$$

- If $y_0 = 1000$ and $a = 10$ then $\Delta U \simeq 0.02$
- If $y_0 = 105000$ and $a = 100000$ then $\Delta U \simeq -2.83$
- In St. Petersburg paradox:
    - Given capital $y_0 = 1000000$ the expected utility increase is positive for $a < 20.88$
    - Given capital $y_0 = 1000$ the expected utility increase is positive for $a < 10.96$
    - Given capital $y_0 = 1$ the expected utility increase is positive for $a < 2.82$

Variance

- Variance of a random variable X is the expected squared deviation from the expectation:

$$D^2(X) = E((X - EX)^2) = \sum_x (x - EX)^2 \cdot P(X = x)$$

- The variance measures the concentration of the distribution around the mean (expectation)
- The square in the definition ensures that deviations with different signs will not cancel each other
- Short formula for variance

$$D^2(X) = E(X^2) - (EX)^2$$

- $D^2(X) \geq 0$
- $D^2(X) = 0$ if and only if X has a degenerate distribution
- $D^2(aX + b) = a^2 D^2(X)$

Distribution	Variance
Bernoulli	$D^2(X) = p(1-p)$
Binomial	$D^2(X) = np(1-p)$
Geometric	$D^2(X) = \frac{1-p}{p^2}$
Negative Binomial	$D^2(X) = \frac{rp}{(1-p)^2}$
Poisson distribution	$D^2(X) = \lambda$

Standard deviation

- The variance has a different “scale” than the random variable itself.
- Standard deviation is defined as the square root of variance:

$$D(X) = \sqrt{D^2(X)}$$

- The standard deviation is often denoted by σ, therefore the variance is denoted by σ^2.

Moments of random variable

- A value:

$$m_k = E(X^k)$$

- is called the k-th moment of a random variable X.
- A value:

$$\mu_k = E((X - EX)^k)$$

- is called the k-th central moment of a random variable X.

Markov’s inequality

- Let X be a nonnegative random variable. For any $a > 0$:

$$P(X \geq a) \leq \frac{EX}{a}$$

Chebyshev's inequality

- For a random variable with finite expected value and finite variance:

$$P(|X - EX| \geq \epsilon) \leq \frac{D^2(X)}{\epsilon^2}$$