Key ideas for topic 9: The Expectation May 27

Independence of random variables Two (discrete) random variables X and Y are indepedendent if

$$\Pr(X = x \cap Y = y) = \Pr(X = x)\Pr(Y = y)$$

for all x, y. In other words observing one doesn't tell anything about the other.

Expectation: The population mean. For a discrete random variable X

$$\mu = \mathbb{E}[X] = \sum_x x\,\mathbf{P}(x),$$

Linearity of expectations. For a collection of random variables $X_1, ..., X_n$

$$\mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i].$$

Expectation of a function $f(s)$ given a random variable X is

$$\mathbb{E}[f(x)] = \sum_x f(x)\mathbf{P}(x),$$

and the distribution of $f(x)$ is

$$\mathbf{P}(f(X)) = \mathbf{P}(f(X) = y) = \sum_{x:g(x)=y} \mathbf{P}(x).$$

Markov's inequality is an example of a law of large numbers, for a positive random variable $X \geq 0$

$$\mathbf{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \quad \forall a > 0.$$

Tail sum formula, for a positive random variable $X \geq 0$

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} \mathbf{P}(X > i).$$

For independent random variables X and Y

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y].$$

Sta-230: Probability **Summer I 2020**

Key ideas for topic 10: Variance and Higher Moments May 28

Moments: The k-th moment of a discrete distribution is

$$M_k = \mathbb{E}(X^k) = \sum_x x^k \, \mathbf{P}(x).$$

Variance: The variance of a random variable measures its spread. For a discrete random variable X

$$\sigma^2 = \mathrm{Var}(X) = \sum_x (x-\mu)^2 \mathbf{P}(x).$$

The standard deviation is $\sigma = \sqrt{\mathrm{Var}(X)}$. Shifting and scaling a r.v. X changes the variance by

$$\mathrm{Var}(aX + b) = a^2 \, \mathrm{Var}(X).$$

If X and Y are independent then

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Given a r.v. X, the random variable $Z = \frac{X-\mu}{\sigma}$ is standardized which means $\mathbb{E}(Z) = 0$ and $\mathrm{Var}(Z) = 1$.

Sums of random variables: Given n random variables $X_1, ..., X_n$ that are independent and identically distributed (iid) the sum

$$S_n = X_1 + \cdots X_n,$$

has mean $n\mu$ and variance $n\sigma^2$. If n is big the the following standardized random variable

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}},$$

is normally distributed with mean 0 and variance 1. This is an example of a central limit theorem. Also

$$\mathbf{P}\left(a \leq \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq b\right) \approx \Phi(b) - \Phi(a),$$

where $\Phi(u)$ is the cumulative distribution function of the standard normal distribution.