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

Key ideas for topic 11: Continuous distributions May 29

Cumulative distribution: The cumulative distribution function (CDF) for a continuous random variable X is

$$F(z) = \mathbf{P}(X \leq z) = \int_{-\infty}^{z} p(x)dx,$$

where $p(z)$ is the probability density function (pdf) for a random variable x

$$p(x) = \lim_{\delta \to 0} \frac{\mathbf{P}\left(x - \delta/2 \leq x \leq x + \delta/2\right)}{\delta}.$$

A standard way of thinking about the pdf is via the cdf

$$p(x) = \frac{dF(x)}{dx}.$$

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

$$\mu = \mathbb{E}[X] = \int_x x\, p(x)df,$$

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

$$\mathbb{E}[f(x)] = \int_x f(x)p(x)dx.$$

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

$$\sigma^2 = \mathrm{Var}(X) = \int_x (x - \mu)^2 p(x)dx.$$

The standard deviation is $\sigma = \sqrt{\mathrm{Var}(X)}$.

The uniform distribution: The uniform distribution is perhaps the most simple continuous distribution. A random variable X follows the uniform distribution on $[0, 1]$, denoted as $X \sim \mathrm{U}(0, 1)$, if X has density

$$f_X(x) = 1,\ 0 < x < 1,$$

and 0 otherwise. In general if $Y = aX + b$, and X as above, then Y follows the uniform distribution $Y \sim \mathrm{U}(b, a + b)$. This can be asserted by looking at the cumulative distributions of X and Y.

Another example of a continuous distribution is the normal distribution that we have seen in binomial approximation.

Key ideas for topic 12: Exponential and Gamma distributions June 01

Exponential distribution: A random variable X follows the exponential distribution with mean $1/\lambda$, $\lambda > 0$ if the density of X is

$$f_X(x) = \lambda e^{-\lambda x},\ x \geq 0,$$

and 0 otherwise. The exponential distribution can be seen as the continuous analog of the geometric distribution. The exponential distribution is a special case of the

Gamma distribution: A random variable X follows the gamma distribution with parameters k, λ, $(k, \lambda > 0)$ if it has the density

$$f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x},\ x \geq 0,$$

and 0 otherwise. The gamma function is a generalization of the factorial $\Gamma(k) = (k-1)!$ for k a positive integer. In that case the Gamma distribution can be seen as the sum of k independent exponential random variables, each with parameter λ.

Both of these distributions can be studied via a **Poisson process** N_t. This is an example of a stochastic process, where for each $t \geq 0$ we assign a Poisson random variable $N_t := N_{t-0} \sim \text{Poisson}(\lambda t)$. In the Poisson process the random variables are connected such that $N_{[t_4 - t_3]}$ and $N_{[t_2 - t_1]}$ are independent whenever $t_1 < t_2 < t_3 < t_4$.

The arrival times T_k of the process follow the Gamma distribution with parameters k, λ, and the inter-arrival times (time between consecutive arrivals) are iid exponential with parameter λ.

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

Key ideas for topic 13: Transformations June 2

Change of variables: For a random variable X with denisty $p_X(x)$ and range (a, b) and $Y = g(X)$ where g is strictly increasing or decreasing on (a, b) with $X = v(Y)$ ($v = g^{-1}$) the density of Y is

$$
\begin{aligned}
p_Y(y) &= p_X(x) \times \left| \frac{dy}{dx} \right|^{-1}, \\
p_Y(y) &= p_X(v(y)) \times v'(y),
\end{aligned}
$$

over $u(a) < y < u(b)$.

This is a consequence of the chain rule applied to the CDFs of X and Y.

Moment generating function (mgf): For X a random variable, the mgf is defined as:

$$\psi_X(t) = \mathbb{E}[e^{tX}],\ t \in (-\infty, \infty).$$

If the mgf is finite in an open interval around $t = 0$, it completely determines the law of the random variable X. For such random variables, The MGF can be used to find moments of X (kth moment is the kth derivative evaluated at $t = 0$) or to find the distribution of sum of iid random variables.