formulas

standard units: z = (original - mean) / standard deviation

β1 = cov(x,y) / var(x)

cov(x,y) = (∑ (y_i - ȳ)(x_i - x̄)) / n

corr(xy) = cov(x,y) / (s_x * s_y)

hierarchical clustering

find successive clusters using previously established clusters
- common approach is bottom-up: start with each element in a separate cluster

single linkage: use minimum distance
complete linkage: use maximum distance
average linkage: use average linkage

"minimum distance between group 1 points and group 2 points is larger than the minimum within-group distance for the same points"

knn classification

k-means

k-means algorithm: 1. construct clusters by associating each point with the closest centroid, 2. calculate new centroids for each set; repeat both till convergence

as k increases, average variance of clusters decreases

use features (x_n) to partition data into K clusters (represented by its **centroid** -- center of points in the cluster)

goal is to minimize intra-class centroid points distances and find cnk (0 - 1 cluster membership) and μk (centroids) that minimize

artificial neural networks

activation functions:
- sigmoid (0, 1)
sigmoid(z) = (exp z) / (1+exp z)
- hyperbolic tangent (-1, 1)
tanh(z) = (exp (z) – exp(-z)) / (exp(z) + exp(-z))