1. For which discrete distribution the mean and the variance are exactly the same? Explain why it's the case.

The Poisson distribution is a discrete probability function that means the variable can only take specific values in a given list of numbers, probably infinite. A Poisson distribution measures how many times an event is likely to occur within “x” period of time. In other words, we can define it as the probability distribution that results from the Poisson experiment. A Poisson experiment is a statistical experiment that classifies the experiment into two categories, such as success or failure. Poisson distribution is a limiting process of the binomial distribution.

For a Poisson Distribution, the mean and the variance are equal. It means that $E(X) = V(X) = \lambda$

The mean and variance of the Poisson distribution
The Poisson distribution was introduced by considering the probability of a single event in a small interval of length h as (λh). We then used the binomial distribution, with $\theta = \lambda h$ and $h = t/n$ and n tending to ∞, to derive the expression for the Poisson distribution. As the mean of the binomial distribution is $n\theta$ it would make sense that the mean of the Poisson distribution is $n\lambda h$. Using $n = t/h$ we get the mean as λt over an interval of length t and therefore the mean is λ over an interval of unit length.
By a similar argument, we know that the variance of the Binomial distribution is $n\theta(1 - \theta)$. Substituting $\theta = \lambda h$ we get the variance as $n\lambda h(1 - \lambda h)$. As n tends to infinity and h to 0 we get the limit λt. Therefore, the variance in unit time is λ.

Reference,

https://byjus.com/maths/poisson-distribution/#:~:text=Are%20the%20mean%20and%20variance,the%20given%20interval%20of%20time.

https://www.sciencedirect.com/topics/mathematics/poisson-distribution

2. If two random variables are independent, their covariance is equal to zero. The converse, however, is not generally true, that is, two variables may have a covariance of zero and still not be independent! Provide examples for two discrete random variables and two continuous random variables (one for each).

$Cov(X, Y) = E((X - \mu X)(Y - \mu Y))$

For continuous random variables,

For an example where the covariance is 0 but X and Y aren't independent, let there be three outcomes, (−1, 1), (0, −2), and (1, 1), all with the same probability 1/3. They're clearly not independent since the value of X determines the value of Y. Note that $\mu X = 0$ and $\mu Y = 0$, so
$Cov(X, Y) = E((X - \mu X)(Y - \mu Y)) = E(XY) = 1/3\,(-1) + 1/3\,(0) + 1/3\,(1) = 0$

For discrete random variables,

There is a simple sample exist for it that y is formula of x,

X is random variable -1 or +1 with probability 0.5

Y is zero for X == -1 and Y is random variable -1 or +1 with probability 0.5 for X == +1

Clearly X and Y are highly dependent (since knowing Y allows me to perfectly know X), but their covariance is zero: They both have zero mean, and

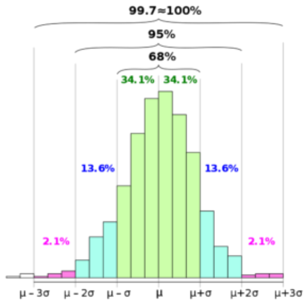
$E(X,Y) = (-1).0\,(P(X=-1)) + 1.1\,(P(X=1,Y=1)) + 1.(-1)\,(P(X=1,Y=1)) = 0$

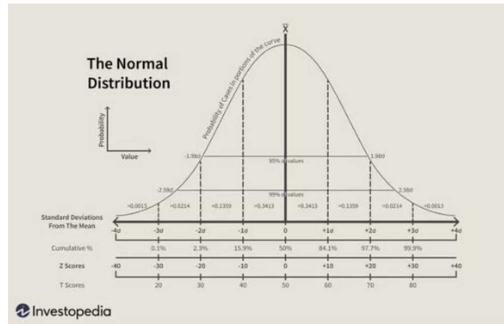
Reference,

https://stats.stackexchange.com/questions/12842/covariance-and-independence

http://aleph0.clarku.edu/~djoyce/ma217/covar.pdf

3. State the Empirical Rule and describe the conditions under which the Empirical Rule may be applied. Give an example.





The empirical rule is also known as the three-sigma rule, as "three-sigma" refers to a statistical distribution of data within three standard deviations from the mean on a normal distribution (bell curve), as indicated by the figure below.

The empirical rule is used often in statistics for forecasting final outcomes.
This probability distribution can thus be used as an interim heuristic since gathering the appropriate data may be time-consuming or even impossible in some cases.
The empirical rule is also used as a rough way to test a distribution's "normality".

The empirical rule is applied to anticipate probable outcomes in a normal distribution.

Let's assume a population of animals in a zoo is known to be normally distributed.

that an animal in the zoo lives to an average of 10 years of age, with a standard deviation of 1.4 years. Assume the zookeeper attempts to figure out the probability of an animal living for more than 7.2 years. This distribution looks as follows:

One standard deviation (μ ± σ): 8.6 to 11.4 years
Two standard deviations (μ ± 2σ): 7.2 to 12.8 years
Three standard deviations ((μ ± 3σ): 5.8 to 14.2 years

The empirical rule states that 95% of the distribution lies within two standard deviations. Thus, 5% lies outside of two standard deviations; half above 12.8 years and a half below 7.2 years. Thus, the probability of living for more than 7.2 years is:

95% + (5% / 2) = 97.5%

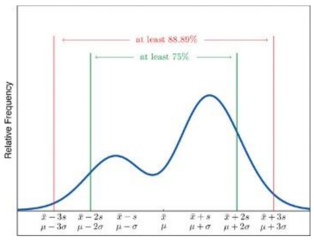
Reference,

https://corporatefinanceinstitute.com/resources/knowledge/other/empirical-rule/

https://www.investopedia.com/terms/e/empirical-rule.asp

4. State Chebyshev’s Theorem and describe the conditions under which Chebyshev’s Theorem may be applied. Give an example.

Chebyshev's Theorem **estimates the minimum proportion of observations that fall within a specified number of standard deviations from the mean**. This theorem applies to a broad range of probability distributions. Chebyshev's Theorem is also known as Chebyshev's Inequality.



The Empirical Rule does not apply to all data sets, only to those that are bell-shaped, and even then is stated in terms of approximations. A result that applies to every data set is known as Chebyshev’s Theorem.

There are two forms of the equation. One determines how close to the mean the data lie and the other calculates how far away from the mean they fall:

Maximum proportion of observations that are more than k standard deviations from the mean	$\frac{1}{k^2}$
Minimum proportion of observations that are within k standard deviations of the mean	$1 - \frac{1}{k^2}$

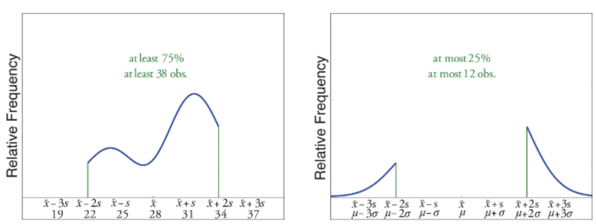
Where k equals the number of standard deviations in which you are interested. K must be greater than 1.

A sample of size n=50 has a mean $\bar{x}$=28 and standard deviation s=3. Without knowing anything else about the sample, what can be said about the number of observations that lie in the interval (22,34)? What can be said about the number of observations that lie outside that interval?

22 < dataset < 34 => 28 – 6 < dataset < 28 + 6 => $\bar{x}$ - 2s < dataset < $\bar{x}$ + 2s => k = 2

¼ of dataset == 50 / 4 = 12.5 (the n should be integer) = 12

¾ of dataset == 3(50) / 4 = 37.5 (the n should be integer) = 38



(a) Within $\bar{x} \pm 2s$ (b) Outside $\bar{x} \pm 2$

Reference,

https://statisticsbyjim.com/basics/chebyshevs-theorem-in-statistics/

The second reference