**Question 1: (Binomial Distribution)**

* That one with n and p
* 2 potential outcomes per trial
* P is the same across all trials.

Suppose the probability of color blinded people = 0.08. we have 10 samples. What is the probability that one man out of 10 having color blindness?

P = 0.08, n = 10.

St.binom.pmf (1, n, p)

What is the probability of having 2 men or less color blinded?

St.binom.cdf (2, n, p)

What is the probability of having 2 men or more color blinded?

P(x>=2) = 1-p(x<=1) = 1 – st.binom.cdf (1, n, p)

**Question 2: (Poisson Distribution)**

* Discrete Distribution
* Describes the number of events occurring in a fixed time interval.
* Requires only lambda that represents expected number of events per time.

A company is advertising on Facebook to direct traffic to the website to sell a product online. The number of click through sales from the ad is Poisson distributed with a mean of 12 click through sales per day.

Find the probability of having:

1. Exactly 10 click through sales in the first day
2. At least 10 click through sales in the first day
3. More than one sale in the first hour
4. =POISSON.DIST(10,12,FALSE) = 0.105
5. P(x>=10) = 1 – P(x<10) = 1 – POISSON.DIST(9,12,TRUE) = 0.758
6. ʎ = 12/24 = 0.5 sales/hour. P(x>1) = 1 – P(x<=1) = 1 – POISSON.DIST(1,0.5,TRUE) = 0.09

**Question 3: (Mean Hypothesis testing and confidence interval)**

A computer manufacturer sets the mean manufacturing cost = 1800$. The company randomly selects 40 computers from its facilities and finds that the mean cost is 1950$ with a standard deviation of 500$. Run a hypothesis test to see if this thought is true. (significance = 5%)

H0 = mean cost <=1800$

H1 = mean cost > 1800$

Test statistic = z = = = 1.897. we have to compare this value with z\_c = st.norm.ppf(0.95,0,1) = 1.6448.

z>z\_c so we have to reject H0.

We can also calculate p\_v = 1 – st.norm.cdf (1.897,0,1) = 0.0289. we compare p\_v with α (0.05). so, p\_v is < 0.05 so again we reject H0.

What is the confidence interval of having 1950 mean and sigma of 500. (α = 0.05)?

Α is divided by 2 because we have 2 tails so α = 0.025 so cdf = 0.975.--> st.norm.ppf (0.975,0,1) = 1.96.

[1950 – 1.96 \*(500/) 1950 + 1.96\*(500/

**Question 4: (chi2 testing):**

**Example one:** 75 students with 11 students out of 75 are left-handed (observed data). The question is does this class fit the theory that 12% of people are left-handed? (One way table, goodness of fit test) (α = 0.05)

|  |  |  |
| --- | --- | --- |
|  | Observed | Expected |
| Left-Handed | 11 | 9 = 12%\*75 |
| Right-Handed | 64 | 66 |
|  | 75 | 75 |

H0 = 12% of people are left-handed.

H1 = the theory is incorrect so different from 12%.

Chi2 = + = 0.505.

DF = 2 – 1 = 1

Chi2\_c = st.chi2.ppf (0.95,1) = 3.841. chi2<chi2\_c🡪H0 accepted. There is not enough evidence to reject H0.

P\_v = 1 – st.chi2.cdf (0.505,1) = 0.477. p\_v > α so again H0 accepted.

**Example two:** 120 people are surveyed for their preferred social media platform. Is there enough evidence to suggest social media preference is independent of sex? α= 0.05.

|  |  |  |
| --- | --- | --- |
|  | Male | Female |
| Facebook | 15 | 20 |
| Insta | 30 | 35 |
| TikTok | 5 | 15 |

|  |  |  |
| --- | --- | --- |
|  | Male | Female |
| Facebook | 14.6 | 20.4 |
| Insta | 27.1 | 37.9 |
| TikTok | 8.3 | 11.7 |

H0 = social media preference is independent of sex.

DF = (r-1)(c-1) = 2.

Chi2 = 2.84.

Chi2\_c = st.chi2.ppf(0.95,2) = 5.99. chi2<chi2\_c so H0 is valid.

P\_v = 1 – st.chi2.cdf(2.84,2) = 0.242. p\_v > α so again H0 is valid.

**Question 5: independent events:**

Given a dice roll, suppose an event A where the roll result is in an even number and an event B where the roll result is equal to or less than 5. A and B are independent.?

A={2,4,6}, B={1,2,3,4,5} respectively, P(A)=1/2, P(B) = 5/6.

𝐴∩𝐵={2,4}

P(𝐴∩𝐵)=1/3

P(A∩B)≠P(A)P(B)

Thus, events A, B are dependent.