S2 - Bible

* Conditions for a binomial distribution
  - A fixed number of trials, n.
  - Each trials should be success or failure
  - The trials are independent
  - The probability of success p, at each trial is constant.

* If $X \sim B(n, p)$ then

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

remember, $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

* If $X \sim B(n, p)$ then

$$E(X) = np$$
$$Var(X) = np(1-p)$$

where p = prob. of success
$(1-p)$ - is prob. of failure

* To find $P(X=x)$ from the table of binomial dist^n function

$$P(X=x) = P(X \leq x) - P(X \leq (x-1))$$

* Poisson distribution
  - If $X \sim Po(\lambda)$ then $P(X=x) = e^{-\lambda} \frac{\lambda^x}{x!}$
  - If $X \sim Po(\lambda) \Rightarrow E(X) = \lambda$ and $Var(X) = \lambda$
  - If $X \sim B(n, p)$ and n is large and p is small then $\Rightarrow X \approx \sim Po(\lambda)$, when $\lambda = np$

Continuous random Variables

- For a continuous random variable, X

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$\sigma^2 = E(X^2) - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

- Cumulative distribution function, $F(x)$
  - $0 \le F(x) \le 1$
  - $F(x) = P(X < x) = \int_{-\infty}^{x} f(t)\,dt$

- The median (m) satisfies $F(m) = 0.5$
- The Lower Quartile Q_1 satisfies $F(Q_1) = 0.25$
- The Upper Quartile Q_3 satisfies $F(Q_3) = 0.75$
- The mode is the x value at the highest point of the f^n.

* Continuous Uniform distribution

- A random variable having a continuous dist^n over the interval (a, b) has a p.d.f

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}$$

- $E(X) = \frac{a+b}{2}$, $\quad Var(X) = \sigma^2 = \frac{(b-a)^2}{12}$

- $$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \le x \le b, \\ 1, & x > b \end{cases}$$

Normal approximations

* The random variable $X \sim B(n, p)$ can be approximated by $Y \sim N(np, np(1-p))$ when:
  - ⊛ n is large
  - ⊛ p is close to 0.5

  A continuous correction should be used

* The random variable $X \sim Po(\lambda)$ can be approximated by $Y \sim N(\lambda, \lambda)$ when λ is large

  A continuous correction should be used.

Populations and Samples

Key words: Census, population, Sample, finite population, infinite population, Sampling unit, Sampling frame, statistic

Hypothesis testing

Key words: hypothesis test, test statistic, critical region, critical values, One-tailed test, two-tailed test.