Shady Haddad

Applied Statistics Formula Sheet 2

<u>Def 3.9)</u>

A random variable *Y* is said to have a *negative binomial probability distribution* if and only if

$$p(y) = \binom{y-1}{r-1} p^r q^{y-r}, \qquad y = r, r+1, r+2, \dots, 0 \le p \le 1.$$

<u>Theorem 3.9</u>

If Y is a random variable with a negative binomial distribution,

$$\mu = E(Y) = \frac{r}{p} \ and \ \sigma^2 = V(Y) = \frac{r(1-p)}{p^2}$$

<u>Def 3.10)</u>

A random variable *Y* is said to have a *hypergeometric probability distribution* if and only if

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

Where *y* is an integer 0, 1, 2, …, *n*, subject to the restrictions $y \le r$ and $n - y \le N - r$.

Theorem 3.10

If Y is a random variable with a hypergeometric distribution,

$$\mu = E(Y) = \frac{nr}{N} \; and \; \sigma^2 = V(Y) = n(\frac{r}{N})(\frac{N-r}{N})(\frac{N-n}{N-1})$$

Def 3.11)

A random variable *Y* is said to have a *Poisson probability distribution* if and only if

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}, \qquad y = 0, 1, 2, \ldots, \lambda > 0$$

Theorem 3.14

Let Y be a random variable with mean μ and finite variance σ^2 . Then, for any constant k > 0,

$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$ or $P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}$

Def 4.1)

Let *Y* denote any random variable. The *distribution function* of *Y*, denoted by *F*(*y*), is such that $F(y) = P(Y \leq y)$ for $-\infty < y < \infty$.

Def 4.3)

Let *F*(*y*) be the distribution function for a continuous random variable *Y*. Then *f(y)*, given by

$$f(y) = \frac{dF(y)}{dy} = F^1(y)$$

Wherever the derivative exists, is called the *probability density function* for the random variable *Y*.

Def 4.5)

The expected value of a continuous random variable *Y* is

$$E(Y) = \int_{-\infty}^{\infty} yf(y)\, dy,$$

Provided that the integral exists.

Theorem 4.4

Let g(Y) be a function of Y; then expected value of g(Y) is given by

$E[g(Y)] = \int_{-\infty}^{\propto} g(y)f(y)dy,$

provided that the integral exists.

Def 4.6)

If $\theta_1 < \theta_2$, a random variable *Y* is said to have a continuous *uniform probability distribution* on the interval (θ_1, θ_2) if and only if the density function of *Y* is

$$f(y) = \begin{cases} \dfrac{1}{\theta_2 - \theta_1}, & \theta_1 \le y \le \theta_2 \\ 0, & elsewhere \end{cases}$$

Def 4.8) (We didn't do normal probability distribution but I included it anyways).

A random variable Y is said to have a normal probability distribution if and only if, for $\sigma > 0 \; and - \infty < \mu < \infty$, the density function of Y is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}$$

Theorem 4.7

If Y is a normally distributed random variable with parameters μ and σ ,then

E(Y)=μ and V(Y)=σ^2.

Def 4.9)

A random variable *Y* is said to have a *gamma distribution with parameters* α > 0 and β > 0 if and only if the density function of *Y* is

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1} e^{\frac{-y}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \le y < \infty, \\ 0, & elsewhere \end{cases}$$

Where

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}\, e^{-y} dy.$$

Theorem 4.8

has a gamma distribution with parameters α and β, then

$\mu = E(Y) = a\beta$ and $\sigma^2 = V(Y) = a\beta^2$

Def 4.10)

Let *v* be a positive integer. A random variable *Y* is said to have a *chi-square distribution with v degrees of freedom* if and only if *Y* is a gamma-distributed random variable with parameters $\alpha = \frac{v}{2}$ and $\beta = 2$.

Def 4.11)

A random variable *Y* is said to have an *exponential distribution with parameter* $\beta > 0$ if and only if the density function of *Y* is

$$f(y) = \begin{cases} \dfrac{1}{\beta} e^{-y/\beta}, & 0 \le y < \infty \\ 0, & elsewhere. \end{cases}$$

Theorem 4.10

is an exponential random variable with parameter β, then

$\mu = E(Y) = \beta$ and $\sigma^2 = V(Y) = \beta^2$

Def 4.12) (We did not do this lesson but included it anyways)

A random variable *Y* is said to have a *beta probability distribution with parameters* $\alpha < 0$ and $\beta > 0$ if and only if the density function of *Y* is

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}(1-y)^{\beta-1}}{\beta(\alpha,\beta)}, & 0 \leq y \leq 1 \\ 0, & elsewhere \end{cases}$$

Where

$$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Theorem 4.13)

Tchebysheff's Theorem Let *Y* be a random variable with finite mean *μ* and variance σ^2. Then, for any *k* > 0,

$$P(|Y - \mu| < k\sigma) \geq 1 - \tfrac{1}{k^2} \text{ or } P(|Y - \mu| \geq k\sigma) \leq \tfrac{1}{k^2}$$

Def 5.1)

Let Y_1 and Y_2 be discrete random variables. The *joint* (or bivariate) *probability function* for Y_1 and Y_2 is given by

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2), \qquad -\infty < y_1 < \infty, -\infty < y_2 < \infty$$

Def 5.2)

For any random variables Y_1 and Y_2, the joint (bivariate) distribution function $F(y_1, y_2)$ is

$$F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2), \qquad -\infty < y_1 < \infty, -\infty < y_2 < \infty.$$

Def 5.3)

Let Y_1 and Y_2 be continuous random variables with joint distribution function $F(y_1, y_2)$. If there exists a nonnegative function $f(y_1, y_2)$, such that

$$F(y_1, y_2) = \int_{-\infty}^{y1} \int_{-\infty}^{y2} f(t_1, t_2) dt_2 dt_1,$$

For all $-\infty < y_1 < \infty, -\infty < y_2 < \infty$, then Y_1 and Y_2 are said to be *jointly continuous random variables*. The function $f(y_1, y_2)$ is called the *joint probability density function*.

Def 5.5)

If Y_1 and Y_2 are jointly discrete random variables with joint probability function $p(y_1, y_2)$ and marginal probability functions $p_1(y_1)$ and $p_2(y_2)$, respectively, then the *conditional discrete probability function* of Y_1 given Y_2 is

$$p(y_1|y_2) = P(Y_1 = y_1|Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} = \frac{p(y_1, y_2)}{p_2(y_2)},$$

Def 5.6)

If Y_1 and Y_2 are jointly continuous random variables with join density function $f(y_1, y_2)$, then the *conditional distribution function* of Y_1 given $Y_2 = y_2$ is

$$F(y_1|y_2) = P(Y_1 \leq y_1|Y_2 = y_2).$$

Def 5.7)

Let Y_1 and Y_2 be jointly continuous random variables with joint density $f(y_1, y_2)$ and marginal densities $f_1(y_1)$ and $f_2(y_2)$, respectively. For any y_2 such that $f_2(y_2) > 0$, the conditional density of Y_1 given $Y_2 = y_2$ is given by

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)}$$

And, for any y_1 such that $f_1(y_1) > 0$, the conditional density of Y_2 given $Y_1 = y_1$ is given by

$$f(y_2|y_1) = \frac{f(y_1, y_2)}{f_1(y_1)}.$$

Def 5.8)

Let Y_1 have distribution function(F_1) (Y_1), Y_2 have distribution function $F_2(y_2)$, and Y_1 and Y_2 have joint distribution function F (Y_1, Y_2). Then Y_1 and Y_2 are said to be independent if and only if

F $(y_1, Y_2) = F_1(Y_1)F_2(Y_2)$

for every pair of real numbers (Y_1, Y_2).

If Y_1 and Y_2 are not independent, they are said to be dependent.