STT 861 Compendium - Part Two

Kenyon Cavender

November 6, 2019

Expected Values

Def The **expected value** of a random variable $g(X)$:

$$\mathbb{E}(g(X)) = \begin{cases} \int_{-\infty}^{\infty} g(x) f_X(x) dx & \text{if } X \text{ is continuous} \\ \sum_X g(x) f_X(x) & \text{if } X \text{ is discrete} \end{cases}$$

Remark If $-\infty \leq \mathbb{E}(g(x)) \leq \infty$ we say the expectation of $g(x)$ exists. Else it does not exist.
notation: $|\mathbb{E}(g(x))| \leq \infty$
In particular, if $g(x) = x$, then we get the expected value of X:

$$\mathbb{E}(g(X)) = \begin{cases} \int x f_X(x) dx & \text{if } X \text{ is continuous} \\ \sum x f_X(x) & \text{if } X \text{ is discrete} \end{cases}$$

This is called the **mean** of r.v. X
Also denoted by μ or μ_X

Theorem Expectation

a. $\mathbb{E}(ag(x) + b) = a\mathbb{E}(g(x)) + ba$,$b$ real constants

b. If$g(x) \geq 0$for all$x \in \mathbb{R}$, then$\mathbb{E}(g(x)) \geq 0$

c. If$g_1(x) \geq g_2(x)$for all$x \in \mathbb{R}$, then$\mathbb{E}(g_1(x)) \geq \mathbb{E}(g_2(x))$

d. For any real constants a,b if $a \leq X \leq b$ then $a < \mathbb{E}(X) < b$

Remark Expectation

a. If g is a linear fn, $\mathbb{E}[g(x)] = g[\mathbb{E}(x)]$

b. $g(x)$ has finite expectation if $0 \leq \mathbb{E}[|g(x)|] \leq \infty$

Moments

Def **Moments** For a r.v. X, we define the r^{th} raw moments by

$$\mu'_r = \begin{cases} \int x' f_X(x) dx & \text{if } X \text{ is continuous} \\ \sum x' f_X(x) & \text{if } X \text{ is discrete} \end{cases}$$

$\mu'_1 = \mathbb{E}(x') = \mathbb{E}(x) = \mu$

Def The r^{th} central moment is defined as $\mu_r = \mathbb{E}[(x - \mu)^r]$

__**Def**__ **Moment Generating Function**: For a r.v. X, the mgf is defined as: $M_X(t) = \mathbb{E}(e^{tx})$ provided the expectation exists for all t in a neighborhood of 0.
I.e. $\mathbb{E} < \infty \,\forall\, t \in (-h, h)$ for some $h > 0$

Theorem 2.3.11a
Let $F_X(x)$ and $F_Y(y)$ be two cdfs all of whose moments exist. If X, Y have bounded support
(i.e. Support$(f_X) = \{x : f_X(x) = 0\}$)
Then $F_X(u) = F_Y(u) \,\forall\, u$ iff $\mathbb{E}(x') = \mathbb{E}(y') \,\forall\, r \in \mathbb{Z}$
Example: P(Heads) for a coin flip and P(Even) for a die roll have the same distribution, but are different variables.

Theorem 2.3.11b
X and Y have identical distributions iff $M_x(t) = M_y(t)$ in some neighborhood of 0.

Theorem 2.3.12
Suppose $\{X_n : n \geq 1\}$ is a sequence of r.v.s with mgf $M_{x_n}(t)$. Then:
$\lim_{n\to\infty} M_{x_n}(t) = M_y(t) \iff \lim_{n\to\infty} F_{x_n}(u) = F_y(u)$ for some r.v. Y for all continuity points u of F_y
X_n converges to Y in distribution

Remark Properties of MGF

a. $M_x(0) = 1$

b. $M_{ax+b}(t) = e^{bt} M_x(at)$

c. $\frac{d^n}{dt^n} M_x(t)|_{t=0} =: M_x^{(n)} = \mu_n'$ (nth raw moment of X)

Theorem Let $f(x)$ be any pdf and let μ and $\sigma > 0$ be any given constants. The the function

$$g(x|\mu, \sigma) = \frac{1}{\sigma} f(\frac{x - \mu}{\sigma})$$

is a pdf.

a. $\mathscr{F} = \{f_0(x - \mu) : \mu \in \mathbb{R}\}$ is called the location family of pdfs and μ is the location parameter

b. $\mathscr{F} = \{\frac{1}{\sigma} f_0(\frac{x}{\sigma}) : \sigma > 0\}$ is called the scale family of pdfs and σ is the scale parameter

c. $\mathscr{F} = \{\frac{1}{\sigma} f_0(\frac{x-\mu}{\sigma}) : \mu \in \mathbb{R},\, \sigma > 0\}$ is called the scale family of pdfs where σ is the scale parameter and μ is the location parameter

Theorem 3.5.6
Let f_0 be a pdf and $\mu \in \mathbb{R}, \sigma > 0$ then a r.v. X has a pdf $f(x) = \frac{1}{\sigma} f_0(\frac{x-\mu}{\sigma})$ iff Z has pdf $f_0(.)$ and $x = \sigma z + \mu$

Theorem Let Z be r.v. w/ pdf $f_0(x)$. Suppose $\mathbb{E}(z) < \infty$ and $Var(z) < \infty$. Then there exists r.v X with pdf $f_0(x) = \frac{1}{\sigma} f_0(\frac{x-\mu}{\sigma})$ which satisfies:

a. $\mathbb{E}(x) = \sigma \mathbb{E}(z) + \mu$

b. $Var(x) = \sigma^2 Var(z)$

__**Def**__ A pdf $f(x|\theta)$ is a member of the exponential family if we can write

$$f(x|\theta) = h(x)c(\theta)exp[\sum_{j=1}^{k} w_j(\theta)t_j(x)]$$

where

a. $h(x) \geq 0$ and does not depend on θ

b. $c(\theta) \geq 0$ and does not depend on x

c. $w_j(\theta)$ doesn't depend on $x \forall j$

d. $t_j(x)$ doesn't depend on $\theta \forall j$

Def

$$\mathbb{I}_A = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Remark Binomial(n,p) where n is also a parameter is not a member of the exponential class. But if n is known, then for the parameter p, binomial is a member of the exponential family.

Theorem 3.4.2
Suppose X has pmf or pdf which belongs to a exponential family. Then

a.

$$\mathbb{E}(\sum_{i=1}^{k} \frac{\delta}{\delta\theta_j} w_i(\theta)t_i(X)) = -\frac{\delta}{\delta\theta_j} logc(\theta)$$

b.

$$\mathbb{E}(\sum_{i=1}^{k} \frac{\delta}{\delta\theta_j} w_i(\theta)t_i(X)) = -\frac{\delta^2}{\delta\theta_j^2} - \mathbb{E}(\sum_{i=1}^{k} \frac{\delta^2}{\delta\theta_j^2} w_i(\theta)t_i(X))logc(\theta)$$

Theorem 3.6.1
Let X be a r.v. and g is a non-negative function. Then for any $r > 0$

$$P(g(x) \geq r) \leq \frac{\mathbb{E}[g(x)]}{r}$$

Def Markov's Inequality
Let X be r.v. then for every $\epsilon > 0$ and $k > 0$

$$P(|x| \geq \epsilon) \leq \frac{\mathbb{E}(|x|^k)}{\epsilon^k}$$

Def Chebyshev's Inequality
Let X be r.v. with mean μ and variance σ^2 Then for any $\epsilon > 0$

$$P(|x - \mu| < \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$