STAT 345/445 Lecture 8

Expected values and moment generating functions – Sections 2.2 and 2.3

Expected values

Definition

Let $X \sim f(x)$. The $\boxed{\textbf{expected value} \text{ or mean.}}$ of $g(X)$ is defined as

- If X is continuous

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

- If X is discrete

$$E(g(X)) = \sum_{x} g(x) f(x)$$

if the integral/sum exists i.e. if $\int_{-\infty}^{\infty} |g(x)| f(x)\, dx < \infty$

In particular the mean of X is :

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx \quad , \quad \text{if cont.}$$

or $E(X) = \sum_{x} x f(x)$, if discrete.

Examples - Find the expected values

- **Exponential distribution** $X \sim \text{Expo}(\beta)$ with $\beta > 0$ and pdf

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & , \; x \geq 0 \\ 0 & , \; x < 0 \end{cases}$$

- **Gamma distribution** $\text{Gamma}(\alpha, \beta)$ with $\alpha > 0$, $\beta > 0$ and pdf

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{(\alpha-1)} e^{-x/\beta} & , \; x \geq 0 \\ 0 & , \; x < 0 \end{cases}$$

Gamma function

Note: $\text{Gamma}(\alpha = 1, \beta) = \text{Expo}(\beta)$, Since $\frac{1}{\Gamma(1)\beta^1} x^{(1-1)} e^{-x/\beta} = \frac{1}{\beta} e^{-x/\beta}$

Let $X \sim \text{Gamma}(\alpha, \beta)$ $\quad E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{\infty} x \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}\,dx$ $\quad \cdots (*)$

Know: $\int_0^{\infty} t^{\alpha-1} e^{-t}\,dt = \Gamma(\alpha)$ and $\int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}\,dx = 1$

Since we know that $f(x)$ is a pdf.

$$* = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \Gamma(\tilde{\alpha})\beta^{\tilde{\alpha}} \underbrace{\underbrace{\int_0^{\infty} \frac{1}{\Gamma(\tilde{\alpha})\beta^{\tilde{\alpha}}} x^{\overset{\tilde{\alpha}}{\alpha+1}-1} e^{-x/\beta}\,dx}_{\text{pdf of Gamma}(\tilde{\alpha}, \beta)}}_{=1} \qquad \boxed{\tilde{\alpha} = \alpha+1}$$

$$= \frac{\Gamma(\tilde{\alpha})\beta^{\tilde{\alpha}}}{\Gamma(\alpha)\beta^{\alpha}} = \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^{\alpha}} = \alpha\beta$$

$$\int_0^\infty x^{\alpha-1} e^{-x/\beta}\, dx = \Gamma(\alpha)\beta^\alpha$$

Some useful facts

- Gamma function $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ for $\alpha > 0$ factorial
  - If n is an integer: $\Gamma(n) = (n-1)!$ $\Gamma(1) = 0! = 1$
  - $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
  - $\Gamma(0.5) = \sqrt{\pi}$
- Integration by parts $\int uv' = uv - \int u'v$
- $e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$ Taylor series for e^λ
- $\int \frac{1}{1+x^2} = \tan^{-1}(x)$

Examples - Find the expected values

- **Poisson distribution** Poisson(λ) with $\lambda > 0$ and pmf

$$f(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & , \ x = 0, 1, 2, \ldots \\ 0 & , \ \text{otherwise} \end{cases}$$

- **Cauchy distribution** with pdf

$$f(x) = \frac{1}{\pi(1 + x^2)} \qquad , \ x \in \mathbb{R}$$

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$$

Let $X \sim \text{Poisson}(\lambda)$

First term is 0.

Substitution $\longleftrightarrow x = t+1$

$$E(X) = \sum_x x f(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=1} \frac{\lambda^x}{(x-1)!} = e^{-\lambda} \sum_{t=0}^{\infty} \frac{\lambda^{t+1}}{t!}$$

Hence. $\displaystyle\sum_{x=0}^{\infty} \frac{(\text{stuff})^x}{x!}$ $\quad x! = x(x-1)(x-2)\cdots 1$

$$\Downarrow$$

$$= e^{-\lambda} \sum_{t=0}^{\infty} \frac{\lambda^{t+1}}{t!}$$

$$= e^{-\lambda} \lambda \sum_{t=0}^{\infty} \frac{\lambda^{t}}{t!}$$

$$= e^{-\lambda} \lambda e^{\lambda} = \lambda$$

Cauchy Expectation does not exist.

Note: $f(x) = \frac{1}{\pi(1+x^2)}$ $x \in \mathbb{R}$, is a pdf:

$$\int_{-\infty}^{\infty} \frac{1}{\pi} \frac{1}{(1+x^2)} dx = \frac{1}{\pi} \tan^{-1}(x) \Big|_{x=-\infty}^{\infty}$$

$$= \frac{1}{\pi}\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = \frac{\pi}{\pi} = 1$$

But $E(X) = \int_{-\infty}^{\infty} x \frac{1}{\pi} \frac{1}{(1+x^2)} dx$ does not exist,

Since $\int_{\infty}^{-\infty} |x| \frac{1}{\pi} \frac{1}{(1+x^2)} dx = \infty$, see book for details

A note on methods

We can approach $E(g(X))$ in two ways:

1. Using the pdf of X:

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

2. Using the pdf of $Y = g(X)$:

$$E(g(X)) = E(Y) = \int_{-\infty}^{\infty} yf_Y(y)dy$$

Properties of expectation

Theorem

Let X be a random variable and let a, b, c be constants, and suppose that $E(g_1(X))$ and $E(g_2(X))$ exist.

(a) $E(ag_1(X) + bg_2(X) + c) = aE(g_1(X)) + bE(g_2(X)) + c$

(b) If $g_1(x) \geq 0$ for all x then $E(g_1(X)) \geq 0$

(c) If $g_1(x) \geq g_2(x)$ for all x then $E(g_1(X)) \geq E(g_2(X))$

(d) If $a \leq g(x) \leq b$ for all x then $a \leq E(g(x)) \leq b$

Expectation is a Linear operator.

$$E(aX+b) = a+bE(X)$$

(a): $E(ag_1(x)+bg_1(x)+c) = \int_{-\infty}^{\infty}(ag_1(x)+bg_1(x)+c)f(x)\,dx$

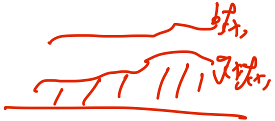
$$= a\int_{-\infty}^{\infty} g_1(x)f(x)\,dx + b\int_{-\infty}^{\infty} g_1(x)f(x)\,dx + c\int_{-\infty}^{\infty} f(x)\,dx$$

$$= aE(g_1(x)) + bE(g_1(x)) + C$$

If $g(x) \le b$, $\forall x \in \mathbb{R}$. $\Rightarrow g(x)f(x) \le bf(x)$

$$E(g(x)) = \int_{-\infty}^{\infty} g(x)\underbrace{f(x)}_{\geq 0}\,dx \le \int_{-\infty}^{\infty} bf(x)\,dx$$

$= b$



Partial proof

The properties on the previous slide are all simple consequences of the definition of $E(g(X))$.

Moments - Section 2.3

Definition

Let n be a positive integer and X be a random variable.

- The *n***th moment** of X is

$$\mu_n' = E\left(X^n\right)$$

- The *n***th central moment** of X is

$$\mu_n = E\left((X-\mu)^n\right)$$

where $\mu = E(X)$

"Cmool: Central moments of our lives" OSU statistics grad students newspaper

Variance

Definition: Mean

The **mean** or **expected value** of a r.v. X is the 1st moment:

$$\mu_1' = \mathrm{E}(X) \equiv \mu$$

Definition: Variance and Standard Deviation

The **variance** of a r.v. X is the 2nd central moment

$$\mathrm{Var}(X) = E\left((X-\mu)^2\right) \equiv \sigma^2$$

The **standard deviation** of X is defined as $\sigma = \sqrt{\mathrm{Var}(X)}$

$$E(X-\mu) = E(X) - \mu = \mu - \mu = 0$$

$$Var(aX+b) = E[(aX+b-a\mu_x-b)^2]$$

Since $E(aX+b) = a\mu_x + b$

$$= E(a^2(X-\mu_x)^2)$$

$$= a^2 E((X-\mu_x)^2) = a^2 Var(X)$$

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

$$= E(X^2 - 2X\mu_x + \mu_x^2) = E(X^2) - 2\mu_x E(X) + \mu_x^2$$

$$= E(X^2) - \mu_X^2 = E(X^2) - E(X)^2$$

$$E(X^2) = \int x^2 f(x)\,dx \qquad V(X) = \int (x-\mu_X)^2 f(x)\,dx$$

$X \sim$ Gamma(α, β) Saw before that $E(X) = \mu_X = \alpha\beta$.

First find $E(X^2) = \int_0^\infty x^2 \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}\,dx$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha}\Gamma(\tilde{\alpha})\beta^{\tilde{\alpha}}\int_0^\infty \frac{1}{\Gamma(\tilde{\alpha})\beta^{\tilde{\alpha}}} x^{\overbrace{\alpha+2}^{=\tilde{\alpha}}-1} e^{-x/\beta}\,dx = \frac{\Gamma(\alpha+2)\beta^{\alpha+2}}{\Gamma(\alpha)\beta^\alpha} = (\alpha+1)\alpha\beta^2$$

$$\Rightarrow \quad Var(X) = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2$$

Possion: $X \sim P(\lambda)$ know $E(X) = \mu_x = \lambda$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} \frac{x e^{-\lambda}\lambda^x}{(x-1)!} \quad \text{let } x = t+1.$$

$$= \sum_{t=0}^{\infty} \frac{(t+1)e^{-\lambda}\lambda^{t+1}}{t!}$$

$$= \lambda \underbrace{\sum_{t=0}^{\infty} \frac{t e^{-\lambda}\lambda^t}{t!}}_{=\lambda} + \lambda \underbrace{\sum_{t=0}^{\infty} \frac{e^{-\lambda}\lambda^t}{t!}}_{=1}$$

$$= \lambda \cdot \lambda + \lambda = \lambda^2 + \lambda$$

$\Rightarrow \quad Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

Properties of Variance

Theorem

Let X be a r.v. with finite variance and let a and b be constants. Then

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

Variance

Theorem

Let X be a r.v. with finite variance. Then

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

Examples

Find the variance for the Exponential, Gamma, and Poisson distributions.

Moment generating functions

- A very useful theoretical tool
  - to characterize a distribution
  - for limits
  - to prove (a version of) the Central Limit Theorem!

Definition

Let $X \sim F(x)$. The **moment generating function (mgf)** of X is defined as

$$M_X(t) = E\left(e^{tX}\right)$$

if the expectation exists for t in a neighborhood of 0.

Moment generating functions

Theorem

If a r.v. X has mgf $M_X(t)$ then

$$E(X^n) = \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}$$

- "Generates" moments

Find the mgf for the Gamma and Poisson distributions and use it to generate the first two moments. $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$ for $x>0$

$X \sim Gamma(\alpha, \beta)$

$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx = \int_0^{\infty} e^{tx} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}\,dx$

$$\int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}+tx} dx \qquad = \frac{-x}{\tilde{\beta}}$$

$$-\frac{x}{\beta}+tx = -x\left(\frac{1-t\beta}{\beta}\right)$$

$$= -\frac{x(1-t\beta)}{\beta}$$

$$= -\frac{x}{\beta/1-t\beta}$$

$$\frac{\beta}{1-t\beta} = \tilde{\beta}$$

$$= -\frac{x}{\tilde{\beta}}$$

$t \neq \frac{1}{\beta}$ ok, if $t < \frac{1}{\beta}$

$$= \frac{1}{\beta^\alpha}\tilde{\beta}^\alpha \int_0^\infty \frac{1}{\Gamma(\alpha)\tilde{\beta}^\alpha} x^{\alpha-1} e^{-x/\tilde{\beta}} dx$$

$$= \frac{\tilde{\beta}^\alpha}{\beta^\alpha} = \frac{\beta^\alpha}{\beta^\alpha(1-t\beta)^\alpha} = \frac{1}{(1-t\beta)^\alpha}$$

$$= (1-\beta t)^{-\alpha}$$

for $t < \frac{1}{\beta}$

Saw before: $E(X) = \alpha\beta$, $E(X^2) = \alpha(\alpha+1)\beta^2$

$$E(X) = \frac{d}{dt}M(t)\Big|_{t=0} = \frac{d}{dt}(1-\beta t)^{-\alpha}\Big|_{t=0} = -\alpha(1-\beta t)^{-(\alpha+1)}(-\beta)\Big|_{t=0}$$

$$= \alpha\beta(1-\beta t)^{-(\alpha+1)}\Big|_{t=0} = \alpha\beta.$$

$$E(X^2) = \frac{d^2}{dt^2}M(t)\Big|_{t=0} = \frac{d}{dt}\alpha\beta(1-\beta t)^{-(\alpha+1)}\Big|_{t=0} = -\alpha\beta(\alpha+1)(1-\beta t)^{-(\alpha+2)}(-\beta)\Big|_{t=0}$$

$$= \alpha\beta^2(\alpha+1)(1-\beta t)^{-(\alpha+2)}\Big|_{t=0} = \alpha(\alpha+1)\beta^2.$$

$X \sim \text{Poisson}(\lambda)$ $\quad f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ $\quad x = 0, 1, 2 \cdots, \lambda > 0$

know: $\displaystyle\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1$ $\quad M(t) = E(e^{tX}) = \displaystyle\sum_{x=0}^{\infty} \frac{e^{tx} e^{-\lambda}\lambda^x}{x!}$

know: $\displaystyle\sum_{x=0}^{\infty} \frac{(\text{stuff})^x}{x!} = e^{\text{stuff}}$, Note $e^{tx} = (e^t)^x$

$$= e^{-\lambda}\sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!} = e^{-\lambda} e^{e^t\lambda} = e^{e^t\lambda - \lambda} = \exp(\lambda\exp(t) - \lambda)$$

$$E(X) = \frac{d}{dt} M(t)\Big|_{t=0} = \frac{d}{dt}\exp(\lambda\exp(t) - \lambda)\Big|_{t=0}$$

$$= \exp(\lambda\exp(t) - \lambda)\cdot\lambda\exp(t)\Big|_{t=0} = \lambda\exp(\lambda\exp(t) - \lambda + t)\Big|_{t=0}$$

$= \lambda e^0 = \lambda$ Same as before

$$E(X^2) = \frac{d}{dt} \lambda \exp(\lambda \exp(t) - \lambda + t) \Big|_{t=0}$$

$$= \lambda \exp(\lambda \exp(t) - \lambda + t)(\lambda \exp(t) + 1) \Big|_{t=0}$$

$$= \lambda \exp(\lambda e^0 - \lambda + 0)(\lambda e^0 + 1) = \lambda e^0(\lambda + 1) = \lambda(\lambda + 1)$$

Same as before

Mgfs uniquely define a distribution

- Mgfs (not moments) uniquely characterize a distribution

Theorem

Let $F_X(x)$ and $F_Y(y)$ be cdfs *for whom all moments exists*.

(a) If X and Y have bounded support, then

$$F_X(u) = F_Y(u)\, \forall u \quad \text{iff} \quad E(X^k) = E(Y^k)\, \forall k = 0, 1, 2, \ldots$$

(b) If mgfs exist and $M_X(t) = M_Y(t)$ for all t in a neighborhood of 0, then $F_X(u) = F_Y(u)$ for all u

- Remember: If $F_X(u) = F_Y(u)\, \forall u$ then $X \stackrel{D}{=} Y$

More on the Theorem

- Note: Moments $E(X^k)$ can exist even when the mgf does not
- Part b): If both $M_X(t)$ and $M_Y(t)$ exist

$$M_X(t) = M_Y(t) \quad \Leftrightarrow \quad X \stackrel{D}{=} Y$$

  - So, just like the cdf and the pdf, the moment generating function (if it exists) uniquely determines the distribution of X

More on the Theorem

- Generally the moments themselves $E(X^k)$ do not uniquely determine a distribution
  - Can have X and Y with same moments for all k but different distribution (and different mgfs)
  - See example 2.3.10 in the textbook

- Part a): If X and Y have *bounded support* we have

$$\text{All moments equal} \quad \Leftrightarrow \quad X \stackrel{D}{=} Y$$

  - So in that special case, the infinite sequence of moments does uniquely determine the distribution

Convergence of mgfs

- Convergence of mgfs implies convergence of cdf's

Theorem

Let $X_1, X_2, X_3, \ldots$ be a sequence of random variables with mgfs $M_{X_i}(t)$, $i = 1, 2, 3, \ldots$ and suppose that

$$\lim_{i\to\infty} M_{X_i}(t) = M_X(t) \quad \forall t \text{ in a neighborhood of } 0$$

and that $M_X(t)$ is an mgf. Then there exists a unique cdf F_X whose moments are determined by $M_X(t)$ and

$$\lim_{i\to\infty} F_{X_i}(x) = F_X(x)$$

Convergence in distribution

for all x where $F_X(x)$ is continuous.

$X_i \xrightarrow{d \text{ distribution}} X$

Poisson approximation to a Binomial

- Let $X_1, X_2, X_3, \ldots$ be a sequence of random variables where

$$X_n \sim \mathrm{Binomial}\left(n, \frac{\lambda}{n}\right) \qquad n \in \mathbb{N}, \lambda > 0$$

- As $n \to \infty$ the distribution of X_n approaches the Poisson distribution.
  - So for large n we can approximate the $\mathrm{Binomial}(n, p)$ distribution with a $\mathrm{Poisson}(np)$ distribution

- proof ... Poisson : mgf : $M_X(t) e^{\lambda e^t - \lambda}$

Binomial(n, p) mgf $M(t) = (pe^t + (1-p))^n$ (see book)

Here $X_n \sim$ Binomial$(n, \frac{\lambda}{n})$, $n \in N$, $\lambda > 0$

$p \Rightarrow \lambda = np$

$$M_{X_n}(t) = \left(\frac{\lambda}{n}e^t + 1 - \frac{\lambda}{n}\right)^n = \left(1 + \frac{\lambda e^t - \lambda}{n}\right)^n \xrightarrow[n\to\infty]{} e^{\lambda e^t - \lambda}$$

$=$ mgf of Poisson(λ)

Set $a_n = \underbrace{\lambda e^t - \lambda}_{\text{no } n \text{ so limit is easy!}}$ $\lim_{n\to\infty} a_n = \lambda e^t - \lambda$

Some useful facts

- **Binomial Theorem:**

$$(x+y)^n = \sum_{i=1}^{n} \binom{n}{i} x^i y^{n-i}$$

for all $x, y \in \mathbb{R}$, $n \in \mathbb{N}$

  - Useful to find the mgf for Binomial distribution - see also textbook example 2.3.9

- **A useful limit.** If $\lim_{n\to\infty} a_n = a$ then

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a$$

More on mgfs

Theorem

Let X be a random variable, a, b constants and $Y = aX + b$. Then

$$M_Y(t) = e^{bt} M_X(at)$$

Proof:

$$M_Y(t) = E(e^{tY}) = E(e^{t(aX+b)}) = E(e^{taX} e^{tb})$$

$$= e^{tb} E(e^{taX}) = e^{tb} M_X(ta)$$

Other special moments

Take a look at exercizes 17 – 19 and 26 - 29 in Ch. 2

- **Mean:** First moment, $\mu = E(X)$
- **Variance:** Second central moment, $\mu_2 = E((X-\mu)^2) = \sigma^2$
- **Skewness:**

$$\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mu_3}{\sigma^3} \qquad \text{where } \mu_3 = E\left((X-\mu)^3\right)$$

making α_3 unitless

  - Measures lack of symmetry
  - A pdf $f(x)$ is **symmetric about a** if

$$f(a-\epsilon) = f(a+\epsilon) \quad \forall \epsilon > 0$$

    - f symmetric $\Leftrightarrow \alpha_3 = 0$
    - f left skewed $\Leftrightarrow \alpha_3 < 0$
    - f right skewed $\Leftrightarrow \alpha_3 > 0$

More special moments and the mode

- **Kurtosis:**

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}$$

– 4th central moment

  - Measures "flatness" versus "peakedness" of $f(x)$

- **Mode** of a distribution is a value a such that $f(a) \geq f(x)$ for all x

Quantiles of a distribution

- If X is a r.v. and $0 < p < 1$ then the value u_p is called the ***p*th quantile** of X if

$$F(u_p) \geq p \qquad \text{and} \qquad 1 - F(u_p) \geq 1 - p$$

- If X is discrete we can define

$$u_p = \min\{x : \; F(x) = p\}$$

- Special cases:
  - **1st quartile** $Q_1 = u_{0.25}$
  - **Median** $Q_2 = m = u_{0.50}$
  - **3rdt quartile** $Q_3 = u_{0.75}$