Chapter 4 Review

Continuous R.V. have "probability density functions" (pdf) $f(x)$ which give probability by integrating (compute area)

- $P(a \le X \le b) = \int_a^b f(x)\,dx$

- $P(X \le b) = \int_{-\infty}^b f(x)\,dx$

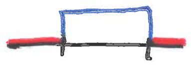
Some pdf are technically not defined for all real x

→ we implicitly set them $= 0$ where not defined

Example: $f(x) = 1/5$ for $1 \le x \le 6$

means

$$f(x) = \begin{cases} 1/5 & \text{for } 1 \le x \le 6 \\ 0 & \text{for other } x \end{cases}$$



Expected value is "center of mass":

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

"average value of $f(x)$" in MAT 119

Variance is the same as before:

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

$$= E[X^2] - (E[X])^2$$

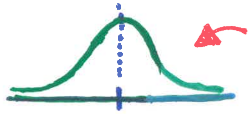
Some Continuous Distributions

- Uniform $[A, B]$ $\qquad f(x) = \frac{1}{B-A}$ for $A \le x \le B$

  X can take any value in interval $[A, B]$

  → All values are equally likely.

- Normal (μ, σ)



symmetric around $x=\mu$

"(Almost) everything is (approximately) normal"

→ Continuous version of Binomial

- $P(|X-\mu| \ge 2\sigma) \approx 5\%$
- $P(|X-\mu| \ge 3\sigma) \approx .3\%$

so we often pretend that Normal(μ, σ) is only defined on $\mu - 3\sigma \le X \le \mu + 3\sigma$

→ Normal$(0, 1)$ $\qquad f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

"Standard Normal" is best normal dist.

Always use "Z" for standard normal.

Z has units: (# std. dev. from mean)

Conversion:

$$\boxed{Z = \frac{X-\mu}{\sigma} \qquad X = \sigma Z + \mu}$$

→ Log Normal

If X cannot be negative, but $\mu - 3\sigma < 0$...

then X cannot be Normal. "Maybe $\ln(X)$ is."

Normal Approximation to Binomial & Poisson

- If n is "big" and p "not too skewed" then

$$\text{Binomial}(n, p) \approx \text{Normal}\left(\mu = np,\ \sigma = \sqrt{np(1-p)}\right)$$

- If λ is "large" then

$$\text{Poisson}(\lambda) \approx \text{Normal}\left(\mu = \lambda,\ \sigma = \sqrt{\lambda}\right)$$

- **Exponential** (λ) $\quad f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$

Waiting time until event occurs $\quad \lambda = \text{rate} = \dfrac{\#\text{occurences}}{\text{length of time}}$

→ Continuous version of Geometric (p)

Divide time into small periods Δt
Assume event cannot occur twice in Δt time period.

$$\#\left(\begin{matrix}\text{time periods until} \\ \text{event occurs}\end{matrix}\right) \sim \text{Geometric}(\lambda \cdot \Delta t)$$

$$\left(\begin{matrix}\text{length of time} \\ \text{until event occurs}\end{matrix}\right) \sim \text{Exponential}(\lambda)$$

Note: Exponential is memoryless

$$P(X \geq x + a \mid X \geq a) = P(X \geq x)$$

"Probability of waiting x more minutes if you've already waited a minutes" = "Probability of waiting x min."

→ Only continuous distribution with this property

Gamma (α, β)

→ Continuous version of Neg. Binomial (r, p)
→ Sum of α Exponential $(1/\beta)$ $\quad$ [Note: $\text{Exp}(\lambda) = \text{Gamma}(1, 1/\lambda)$]

$$\left(\begin{matrix}\text{length of time until} \\ \text{event occurs } \alpha \text{ times}\end{matrix}\right) \sim \text{Gamma}(\alpha, \beta)$$

$$\beta = 1/\lambda = \frac{\text{length of time}}{\#\text{occurences}}$$

If $\alpha = n$ is an integer, then this is sometimes called the "Erlang Distribution" $f(x) = \dfrac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$

"Gamma" Distribution because changing $n \longmapsto \alpha$ requires generalizing factorial: $(n-1)! \longmapsto \Gamma(\alpha)$ "Gamma" funct

(Life expectancy is usually considered to be Gamma
↳ "After α things fail, you die...")

- **Sum law for Gamma:**

If $\begin{cases} X_1 \sim \text{Gamma}(\alpha_1, \beta) \\ X_2 \sim \text{Gamma}(\alpha_2, \beta) \end{cases}$ "time for α_1 occurrences" / "time for α_2 occurrences"

then $(X_1 + X_2) \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$ "time for $\alpha_1 + \alpha_2$ occurrences"

- **Scaling law for Gamma:**

If $X \sim \text{Gamma}(\alpha, \beta)$ then $X/\beta \sim \text{Gamma}(\alpha, 1)$

"rate $= \dfrac{1 \text{ min}}{1 \text{ occurence}}$"

Gamma and Other Distributions

Exponential

$$\text{Exponential}(\lambda) = \text{Gamma}(1, 1/\lambda)$$

[If $X_1, \dots, X_n$ are all Exponential (λ)
then $(X_1 + \dots + X_n) \sim \text{Gamma}(n, 1/\lambda)$]

Normal

If $X \sim \text{Normal}(0, \sigma)$
then $X^2 \sim \text{Gamma}(1/2, 2\sigma)$

If $X_1, \dots, X_n$ are all Normal $(0, \sigma)$
then $(X_1^2 + \dots + X_n^2) \sim \text{Gamma}(n/2, 2\sigma)$

In particular

$$\left(X_1^2/\sigma^2 + \dots + X_n^2/\sigma^2\right) \sim \text{Gamma}(n/2, 2)$$

$$\parallel$$

$$\left(\frac{X_1^2 + \dots + X_n^2}{\sigma^2}\right)$$

[This is also called $\chi^2(n)$ "Chi-Squared with n degrees of freedom"]

Using R with continuous distributions.

Plot pdf using "curve" command

Example: $X \sim \text{Normal}(0, 1)$ ‖ $X \sim \text{Normal}(2, 3)$

curve (dnorm, -3, 3) ‖ curve (dnorm (x, 2, 3), -7, 11)

normal pdf

begin/end x-values for plot

Example: $X \sim \text{Exponential}(1/3)$

curve (dexp (x, 1/3), 0, 20)

Example: $X \sim \text{Gamma}(2, 3)$

curve (dgamma (x, 2, 3), 0, 5)

Evaluate probabilities the same as before.

Example: $X \sim \text{Normal}(2, 3)$

Compute $P(X < 5)$.
pnorm (5, 2, 3)

Compute $P(X > 5)$.
1 - pnorm (5, 2, 3)

Compute $P(1 < X < 5)$.
pnorm (5, 2, 3) — pnorm (1, 2, 3)

Find critical x so that $P(X < x) = .05$
qnorm (.05, 2, 3)