ST2334 Midterms Cheatsheet v1.0 (2019-05-01)
by Julius Putra Tanu Setiaji, page 1 of 2

1 Basic Concepts of Probability

- **Sample Space** (S) = set of all possible outcomes of a statistical experiment
- **Sample Points** = An element of the sample space
- **Event** = Subset of a sample space
- Sample space = **sure event**, subset of $S = \varnothing$ = **null event**
- **Mutually exclusive/disjoint** if $A \cap B = \varnothing$
- **Contained**: $A \subset B \equiv B \supset A$.
- If $A \subset B$ and $B \supset A$, then $A = B$

1.1 Basic Properties

1. $A \cap A' = \varnothing$
2. $A \cap \varnothing = \varnothing$
3. $A \cup A' = S$
4. $(A')' = A$
5. $(A \cap B)' = A' \cup B'$
6. $(A \cup B)' = A' \cap B'$
7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
8. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
9. $A \cup B = A \cup (B \cap A')$
10. $A = (A \cap B) \cup (A \cap B')$

1.2 De Morgan's Law

1. $(\bigcup_{r=1}^{n} A_r)' = \bigcap_{r=1}^{n} (A_r)'$
2. $(\bigcap_{r=1}^{n} A_r)' = \bigcup_{r=1}^{n} (A_r)'$

1.3 Counting Methods

1.3.1 Multiplication & Addition Principle

1.3.2 Permutation

- An arrangement of r objects from a set of n objects, $r \le n$, order taken into consideration.
- n distinct objects taken r at a time $= {}_nP_r = \frac{n!}{(n-r)!}$
- In a circle: $(n-1)!$
- Not all are distinct: $\sum_{r=1}^{k} n_k = n$, ${}_nP_{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1!n_2!\ldots n_k!}$

1.3.3 Combination

- No of ways selecting r from n objects w/o regarding order
- $\binom{n}{r} = {}_nC_r = \frac{n!}{r!(n-r)!}$, ${}_nC_r \times r! = {}_nP_r$
- $\binom{n}{r}$ = binomial coefficient of the term $a^r b^{n-r}$ in binomial expansion of $(a+b)^n$:
  1. $\binom{n}{r} = \binom{n}{n-r}$ for $r = 0, 1, \ldots, n$
  2. $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$ for $1 \le r \le n$
  3. $\binom{n}{r} = 0$ for $r < 0$ pr $r > n$

1.4 Relative frequency (f_A)

$f_A = \frac{n_A}{n}$, event A in n repetitions of experiment E, n_A = no of times that event A occured among the n repetitions.

1.4.1 Properties

1. $0 \le f_A \le 1$
2. $f_A = 1$ iff A occurs every time among the n repetitions
3. $f_A = 0$ off A never occurs among the n repetitions
4. Events A and B are **mutually exclusive** $\rightarrow f_{A \cup B} = f_A + f_B$
5. f_A "stabilises" near some definite numerical value as the experiment is repeated more and more times.

1.5 Axioms of Probability

1. $0 \le \Pr(A) \le 1$
2. $\Pr(S) = 1$
3. If $A_1, A_2, \ldots$ are mutually exclusive (disjoint), i.e. $A_i \cap A_j = \varnothing$ when $i \ne j$, then $\Pr(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \Pr(A_i)$
   In particular, if events A and B are mutually exclusive, then $\Pr(A \cup B) = \Pr(A) + \Pr(B)$

1.6 Properties of Probability

1. $\Pr(\varnothing) = 0$
2. If $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then $\Pr(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \Pr(A_i)$
3. $\Pr(A') = 1 - \Pr(A)$
4. $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B')$
5. $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
6. $\Pr(A \cup B \cup C) = \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(B \cap C) - \Pr(A \cap C) + \Pr(A \cap B \cap C)$
7. The Inclusion-Exclusion Principle
   $\Pr(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \Pr(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Pr(A_i \cap A_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \Pr(A_i \cap A_j \cap A_k) - \ldots$
8. If $A \subset B$, then $\Pr(A) \le \Pr(B)$

1.7 Conditional Probability, $P(A \mid B)$

- $\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$, if $\Pr(A) \ne 0$
- For fixed A, $\Pr(B \mid A)$ satisfies the postulates of probability.
- False positive: $\Pr(+ \mid \text{condition})$

1.7.1 Multiplication rule

- $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A) = \Pr(B)\Pr(A \mid B)$, providing $\Pr(A) > 0, \Pr(B) > 0$
- $\Pr(A \cap B \cap C) = \Pr(A)\Pr(B \mid A)\Pr(C \mid A \cap B)$
- $\Pr(A_1 \cap \ldots \cap A_n) = \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1 \cap A_2) \ldots \Pr(A_n \mid A_1 \cap \ldots \cap A_{n-1})$

1.7.2 The Law of Total Probability

- Let $A_1, A_2, \ldots, A_n$ be a partition of sample space S (mutually exclusive and exhaustive events s.t. $A_i \cap A_j = \varnothing$ for $i \ne j$ and $\cup_{i=1}^{n} A_i = S$).
- Then $\Pr(B) = \sum_{i=1}^{n} \Pr(B \cap A_i) = \sum_{i=1}^{n} \Pr(A_i)\Pr(B \mid A_i)$

1.7.3 Bayes' Theorem

- Let $A_1, A_2, \ldots, A_n$ be a partition of S
- $\Pr(A_k \mid B) = \frac{\Pr(A_k)\Pr(B \mid A_k)}{\sum_{i=1}^{n} \Pr(A_i)\Pr(B \mid A_i)} = \frac{\Pr(A_k)\Pr(B \mid A_k)}{\Pr(B)}$, $k \in [1, n]$

1.8 Independent Events

- Definition: iff $\Pr(A \cap B) = \Pr(A)\Pr(B)$

1.8.1 Properties

- Suppose $\Pr(A) > 0, \Pr(B) > 0$, A and B are independent:
  - $\Pr(B \mid A) = \Pr(B)$ and $\Pr(A \mid B) = \Pr(A)$
  - A and B cannot be mutually exclusive (and vice versa)
- The sample space S and $\varnothing$ are independent of any event
- If $A \subset B$, then A and B are dependent unless $B = S$

Warning: Indep events can't be shown using Venn Diagram, hence calc! Cannot use intuition

1.8.2 Theorem

If A, B are indep, then so are A and B', A' and B, A' and B'.

1.8.3 n Independent Events

- **Pairwise Independent Events**:
  Events $A_1, A_2, \ldots, A_n$ are pairwise indep iff $\Pr(A_i \cap A_j) = \Pr(A_i)\Pr(A_j)$
- **Mutually Independent**:
  Events $A_1, A_2, \ldots, A_n$ are (mutually) independent iff for any subset $\{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ of $A_1, A_2, \ldots, A_n$,
  $\Pr(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \Pr(A_{i_1})\Pr(A_{i_2}) \ldots \Pr(A_{i_k})$

1.8.4 Remarks

- $A_1, A_2, \ldots, A_n$ are mutually independent $\Leftrightarrow$ for any pair of events A_j, A_k where $j \ne k$, the multiplication rule holds, for any 3 distinct events, the multiplication rule holds, and so on $\Pr(A_1 \cap A_2 \cap \ldots \cap A_n) = \Pr(A_1)\Pr(A_2) \ldots \Pr(A_n)$.
  In total there are $2^n - n - 1$ different cases.
- Mutually indep $\Rightarrow$ pairwise indep (not the converse)
- Suppose $A_1, A_2, \ldots, A_n$ are mutually indep events, let $B_i = A_i$ or A'_i, $i \in [1, n]$. Then $B_1, B_2, \ldots, B_n$ are also mutually indep events.

2 Concepts of Random Variables

2.1 Random Variable

2.1.1 Definition

Let S be sample space assoc with experiment E. R.V. is a function X, which assigns a number to every element $s \in S$

2.1.2 Notes

- X is a real-valued function
- Range space of X, $R_X = \{x \mid x = X(s), s \in S\}$.
- Each possible value x of X represents an event that is a subset of the sample space S
- If S has elements that are themselves real numbers, we take $X(s) = s$. In this case $R_X = S$

2.2 Equivalent Events

2.2.1 Definition

- Let E be an experiment in sample space S. Let X be an R.V. defined on S, and R_x its range space, i.e. $X : S \rightarrow \mathbb{R}$
- Let B be an event w.r.t. R_X, i.e. $B \subset R_X$
- Suppose $A = \{s \in S \mid X(s) \in B\}$
  (A consists of all sample points s in S for which $X(s) \in B$)
- A and B are **equivalent events**, and $\Pr(B) = \Pr(A)$

2.2.2 Example

- Consider tossing a coin twice, $S = \{HH, HT, TH, TT\}$
- Let X be no of heads, then $R_X = \{0, 1, 2\}$
- $A_1 = \{HH\}$ equiv $B_1 = \{2\}$, $A_2 = \{HT, TH\}$ equiv $B_2 = \{1\}$, $A_3 = \{TT\}$ equiv $B_3 = \{0\}$, $A_4 = \{HH, HT, TH\}$ equiv $B_4 = \{2, 1\}$

2.3 Discrete Probability Distributions

2.3.1 Discrete R.V.

Let X be an R.V. If R_X is **finite or countable infinite**, X is discrete R.V.

2.3.2 Probability Function (p.f.) or Probability Mass Function (p.m.f.)

- For a discrete R.V., each value X has a certain probability $f(x)$. Such a function $f(x)$ is called the p.f.
- The collection of pairs $(x_i, f(x_i))$ is prob distribution of X
- The probability of $X = x_i$ denoted by $f(x_i)$ must satisfy:
  1. $f(x_i) \ge 0 \; \forall \, x_i$
  2. $\sum_{i=1}^{\infty} f(x_i) = 1$

2.4 Continuous Probability Distributions

2.4.1 Continuous R.V.

Suppose that R_X is an **interval or a collection of intervals**, then X is a continuous R.V.

2.4.2 Probability Density Function (p.d.f.)

- Let X be a continuous R.V.
- p.d.f. $f(x)$ is a function satisfying:
  1. $f(x) \ge 0 \; \forall \, x \in R_X$
  2. $\int_{R_X} f(x)\,dx = 1$ or $\int_{-\infty}^{\infty} f(x)\,dx = 1$ as $f(x) = 0 \; \forall \, x \notin R_X$
  3. $\forall \, c, d : c < d$ (i.e. $(c, d) \subset R_X$), $\Pr(c \le X \le d) = \int_c^d f(x)\,dx$

2.4.3 Remarks

- $\Pr(c \le X \le d) = \int_c^d f(x)\,dx$ represents area under the graph of the p.d.f. $f(x)$ between $x = c$ and $x = d$
- Let x_0 be a fixed value, $\Pr(X = x_0) = 0$
- $\le$ and $<$ can be used interchangeably in a prob statement.
- $\Pr(A) = 0$ does not necessarily imply $A = \varnothing$
- $R_X \in [a, b] \Rightarrow f(x) = 0 \; \forall \, x \notin [a, b]$

2.5 Cumulative Distribution Function (c.d.f.)

- Let X be an R.V., discrete or continuous.
- $F(x)$ is a c.d.f. of X where $F(x) = \Pr(X \le x)$

2.5.1 c.d.f. for Discrete R.V.

- $F(x) = \sum_{t \le x} f(t) = \sum_{t \le x} \Pr(X = t)$
- c.d.f. of a discrete R.V. is a step function
- $\forall \, a, b$ s.t. $a \le b$, $\Pr(a \le X \le b) = \Pr(X \le b) - \Pr(X < a) = F(b) - F(a^-)$ where a^- is the largest possible value of X that is strictly less than a
- $R_X \subset \mathbb{Z}, a, b \in \mathbb{Z} \Rightarrow$
  - $\Pr(a \le X \le b) = \Pr(X = a \text{ or } a+1 \text{ or } \ldots \text{ or } b) = F(b) - F(a-1)$
  - Taking $a = b$, $\Pr(X = a) = F(a) - F(a-1)$

2.5.2 c.d.f. for Continuous R.V.

- $F(x) = \int_{-\infty}^{x} f(t)\,dt$
- $f(x) = \frac{dF(x)}{dx}$ if the derivative exists
- $\Pr(a \le X \le b) = \Pr(a < X \le b) = F(b) - F(a)$
- $F(x)$ is a non-decreasing function: $x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$
- $0 \le F(x) \le 1$

2.6 Mean and Variance of an R.V.

2.6.1 Expected Value / Mean / Mathematical Expectation

- **Discrete:** $E(X) = \mu_X = \sum_i x_i f(x_i) = \sum_x x f(x)$
- If $f(x) = \frac{1}{N}$ for each of the N values of x, $E(X) = \frac{1}{N} \sum_i x_i$
- **Continuous:** $E(X) = \mu_X = \int_{-\infty}^{\infty} x f(x)\,dx$
- **Remark:** The expected value exists provided the sum/integral exists

2.6.2 Expectation of a function of an R.V.

$\forall \, g(X)$ with p.f. $f_X(x)$

- **Discrete:** $E[g(X)] = \sum_x g(x) f_X(x)$
- **Continuous:** $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$
- Provided the sum/integral exists.

2.6.3 Variance ($\sigma_X^2 = V(X)$)

- $g(x) = (x - \mu_X)^2$, Let X be an R.V. with p.f. $f(x)$
- $\sigma_X^2 = V(X) = E[(X - \mu_X)^2]$
- $E[(X - \mu_X)^2] = \begin{cases} \sum_x (x - \mu_X)^2 f_X(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\,dx & \text{if } X \text{ is continuous} \end{cases}$
- $V(X) \ge 0$, $V(X) = E(X^2) - [E(X)]^2$
- **Standard deviation** $= \sigma_X = \sqrt{V(X)}$

2.6.4 K-th moment of X

- **Definition:** $E(X^k)$, use $g(x) = x^k$ in expectation of a fn

2.6.5 Properties of Expectation

1. $E(aX + b) = aE(X) + b$
2. $V(X) = E(X^2) - [E(X)]^2$
3. $V(aX + b) = a^2 V(X)$

2.7 Chebyshev's Inequality

- Let X be an R.V. with $E(X) = \mu$, $V(X) = \sigma^2$
- $\forall \, k > 0$, $\Pr(|X - \mu| > k\sigma) \le \frac{1}{k^2}$
- Alternatively, $\Pr(|X - \mu| \le k\sigma) \ge 1 - \frac{1}{k^2}$
- Holds for **all** distributions with finite mean and variance
- Gives a **lower bound** but not exact probability.

3 MA1521 Shit

3.1 Taylor Series of f at a

$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x - a)^k$

3.2 MacLaurin Series

Taylor series of f at 0, i.e. $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$

3.3 List of common MacLaurin Series

- $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, -1 < x < 1, R = 1$
- $\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n, -1 < x < 1, R = 1$
- $\frac{1}{1+x^2} = \sum_{n=0}^{\infty} x^2 n, -1 < x < 1, R = 1$
- $ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n}, -1 < x < 1, R = 1$
- $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, -\infty < x < \infty, R = \infty$
- $\cos x = \sum_{n=1}^{\infty} \frac{(-1)^n x^2 n}{(2n)!}, -\infty < x < \infty, R = \infty$
- $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, -\infty < x < \infty, R = \infty$
- $\tan^{-1} x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}, -1 \le x \le 1, R = 1$
- $\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1}, -1 < x < 1, R = 1$
- $\frac{1}{(1-x)^3} = \frac{1}{2} \sum_{n=2}^{\infty} n(n-1) x^{n-2}, -1 < x < 1, R = 1$
- $(1+x)^k = \sum_{n=0}^{\infty} \binom{k}{n} x^n, -1 < x < 1, R = 1$
- $(1+x)^n = 1 + nx + \frac{n(n-1)}{2!} x^2 + \frac{n(n-1)(n-2)}{3!} x^3 + \ldots, -1 < x < 1, R = 1$

3.4 Indefinite Integral

Denoted by $\int f(x)dx = F(x) + C$

3.5 Rules of Indefinite Integration

1. $\int kf(x)dx = k\int f(x)dx$
2. $\int -f(x)dx = -\int f(x)dx$
3. $\int [f(x) \pm g(x)]dx = \int f(x)dx \pm \int g(x)dx$

3.6 Integral Formulae

Function	**Integral**
$\int \cot x dx$	$\ln(\sin x) + C$
$\int \sec x \tan x dx$	$\sec x + C$
$\int \csc x \cot x dx$	$\csc x + C$
$\int \sec^2 x dx$	$\tan x + C$
$\int \csc^2 x dx$	$-\cot x + C$
$\int x^n dx$	$\frac{x^{n+1}}{n+1} + C, n \neq -1, n$ rational
$\int \frac{1}{\sqrt{a^2-x^2}} dx$	$\sin^{-1}(\frac{x}{a}) + C$
$\int \frac{1}{a^2+x^2} dx$	$\frac{1}{a}\tan^{-1}(\frac{x}{a}) + C$
$\int 1 dx = \int dx$	$x + C$
$\int e^x dx$	$e^x + C$
$\int a^x dx$	$\frac{a^x}{\ln a}$
$\int \ln x dx$	$x \ln x - x + C$
$\int \frac{1}{x} dx$	$\ln x + C$
$\int \sin kx dx$	$-\frac{\cos kx}{k} + C$
$\int \cos kx dx$	$\frac{\sin kx}{k} + C$
$\int \tan^2 x dx$	$\tan x - x + C$
$\int \sec x dx$	$\ln(\sec x + \tan x) + C$
$\int \csc x dx$	$\ln(\csc x - \cot x) + C$

3.7 Riemann (Definite) Integrals

Riemann sum on f on $[a,b] \approx \sum_{k=1}^{n} f(c_k)\Delta x$

Exact area = $\lim_{n\to\infty} \sum_{k=1}^{n} f(c_k)\Delta x$

Riemann Integral of f over $[a,b]$:

$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{k=1}^{n} f(c_k)\Delta x$

3.8 Rules of Definite Integrals

1. $\int_a^a f(x)dx = 0$, $\int_a^b kf(x)dx = k\int_a^b f(x)dx$
2. $\int_a^b f(x)dx = -\int_b^a f(x)dx$
3. $\int_a^b [f(x) \pm g(x)] = \int_a^b f(x) \pm \int_a^b g(x)$
4. If $f(x) \geq g(x)$ on $[a,b]$, then $\int_a^b f(x)dc \geq \int_a^b g(x)dx$
   If $f(x) \geq 0$ on $[a,b]$, then $\int_a^b f(x)dx \geq 0$
5. If f is continuous on the interval joining a, b and c, then $\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$

3.9 Fundamental Thm of Calculus

$F'(x) = f(x)$ If F is an antiderivative of f on $[a,b]$, then $\int_a^b F'(x)dx = \int_a^b f(x)dx = F(b) - F(a)$

x‘ Let f be continuous on $[a,b]$. Then

$\frac{d}{dx}\int_a^x f(t)dt = f(x)$

Note the 2 x's: on $\frac{d}{dx}$ and $\int_a^x$ and $f(t)$ is indep of x

1. $\frac{d}{dx}\int_0^2 t^2 dt = 0$, $\frac{d}{dx}\int_0^x \sin\sqrt{t}dt = \sin\sqrt{x}$
2. $\frac{d}{dx}\left(\int_1^{x^4} \frac{t}{\sqrt{t^3+2}}dt\right) = \frac{d}{dx^4}\left(\int_1^{x^4} \frac{t}{\sqrt{t^3+2}}dt\right)\frac{dx^4}{dx}$
   $= \frac{x^4}{\sqrt{(x^4)^3+2}}(4x^3) = \frac{4x^7}{\sqrt{x^1 2+2}}$
3. $\frac{d}{dx}\int_x^a f(t)dt = -\frac{d}{dx}\int_a^x f(t)dt$
4. $\frac{d}{dx}\int_{x^2}^{x^4} f(t)dt = \frac{d}{dx}\int_a^{x^4} f(t)dt - \frac{d}{dx}\int_a^{x^2} f(t)dt$

3.10 Integration Methods

• **Integration by Substitution** :

Use the form $\int f(g(x))dg(x)$ OR use a dummy variable to get to a form in the Integral Formulae (taking into account chain rule)

Integral	**Sub**	**Use identity**
$a^2 - u^2$	$u = a\sin\theta$	$1 - \sin^2\theta = \cos^2\theta$
$a^2 + u^2$	$u = a\tan\theta$	$1 + \tan^2\theta = \sec^2\theta$
$u^2 - a^2$	$u = a\sec\theta$	$sec^2\theta - 1 = \tan^2\theta$

• **Integration by Part** :

$\int uv'dx = uv - \int u'vdx$

Choose u by LIATE (Logarithmic, Inverse trigo, Algebraic, Trigo, Exponential)

3.11 Derivative Formulae

Function	Derivative
$(f(x))^n$	$nf'(x)f(x)^{n-1}$
$\sin f(x)$	$f'(x)\cos f(x)$
$\cos f(x)$	$-f'(x)\sin f(x)$
$\tan f(x)$	$f'(x)\sec^2 f(x)$
$\cot f(x)$	$-f'(x)\csc^2 f(x)$
$\sec f(x)$	$f'(x)\sec f(x)\tan f(x)$
$\csc f(x)$	$-f'(x)\csc f(x)\cot f(x)$
$a^f(x)$	$f'(x)a^{f(x)}\ln a$
k	0
$e^f(x)$	$f'(x)e^{f(x)}$
$\log_a f(x)$	$\frac{f'(x)}{f(x)\ln a}$
$\ln f(x)$	$\frac{f'(x)}{f(x)}$
$\sin^{-1} f(x)$	$\frac{f'(x)}{\sqrt{1-f(x)^2}}$
$\cos^{-1} f(x)$	$-\frac{f'(x)}{\sqrt{1-f(x)^2}}$
$\tan^{-1} f(x)$	$\frac{f'(x)}{1+f(x)^2}$

3.12 Rules of Differentiation

- $(kf)'(x) = kf'(x)$
- $(f \pm g)'(x) = f'(x) \pm g'(x)$
- $\frac{d}{dx}uv = u\frac{dv}{dx} + v\frac{du}{dx}$
- $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$
- $\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$ or $\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx}$