# MATH323 — PROBABILITY
# FINALS REVIEW

## MOMENT GENERATING FUNCTION

The $k^{th}$ moment of a r.v. $Y$ taken about the origin: $E(Y^k) = \mu'_k$

The $k^{th}$ moment taken about the mean is the $k^{th}$ central moment

$\mu^*_k = E((Y-\mu)^k)$.

Moment generating function: $m(t) = E(e^{tY})$ ← for discrete: $\sum_y e^{ty} p(y)$; for cont: $\int_{-\infty}^{\infty} e^{ty} f(y)\, dy$

- $m(t)$ exists if $\exists\, b > 0$ s.t. $m(t)$ is finite for $|t| \le b$

- Correspondance Theorem: There is a one-to-one correspondance between the mgf and pdf. ie The mgf for each probability distribution is unique. If $Y$, $Z$, 2 r.v.s, have the same mgf then they have the same probability distribution too.
- If we can find $E(e^{tY})$, we can find any moment of $Y$.

  If $m(t)$ exists, then for any $k \in \mathbb{Z}^+$:

$$\left.\frac{d^k m(t)}{dt^k}\right|_{t=0} = m^{(k)}(0) = \mu'_k = E(Y^k)$$

* mgfs are included in formula sheet!

## TCHEBYSHEFF'S THEOREM

Let $Y$ be an r.v. with mean $\mu$ and finite variance $\sigma^2$. Then for any $k > 0$,

not on formula sheet! ↳ $P(|Y-\mu| < k\sigma) \ge 1 - \frac{1}{k^2}$ or $P(|Y-\mu| \ge k\sigma) \le \frac{1}{k^2}$

This theorem allows us to approximate certain probabilities when only $\mu$ and $\sigma^2$ are known.

eg ~75% of the data falls within $2\sigma$ of the mean,
~89% of the data falls within $3\sigma$ of the mean

## TRANSFORMING A VARIABLE

$Y$ a continuous r.v. with pdf $f(y)$, $U = h(y)$

1. Find $F_U(u) = P(U \le u) = P(h(y) \le u)$
   $= P(\cancel{h}(y) \le h^{-1}(u))$

   by integrating $f(y)$ over the region for $U \le u$.
2. $f(u) = \frac{d}{du} F(u) = \frac{d}{du} F_Y(h^{-1}(u))$

Hilroy

eg $Y$ cont r.v., $f(y) = \begin{cases} 2(1-y), & 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$

let $U = 1 - 2Y$, find $f_U(u)$.

$u = 1 - 2y \Rightarrow y = \frac{1-u}{2}$

$\cancel{f(u) = 2\left(1 - \left(\frac{1-u}{2}\right)\right) \cdot \frac{d}{du}\left(h^{-1}(u)\right)}$

$$
\begin{aligned}
f_U(u) &= f_Y\left(h^{-1}(u)\right) \cdot \left|\frac{d}{du} h^{-1}(u)\right| \\
&= f_Y\left(\frac{1-u}{2}\right) \cdot \left|-\frac{1}{2}\right| \\
&= 2\left(1 - \left(\frac{1-u}{2}\right)\right)\left(\frac{1}{2}\right) \\
&= \frac{1+u}{2}
\end{aligned}
$$

$1 - 2(0) = 1$
$1 - 2(1) = -1$

$f_U(u) = \begin{cases} \frac{1+u}{2}, & -1 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$

In general, $f_U(u) = f_Y\left(h^{-1}(u)\right) \cdot \left|\frac{d}{du} h^{-1}(u)\right|$

## MULTIVARIATE DISTS.

let $Y_1$ and $Y_2$ be $\boxed{\text{discrete}}$ r.v.s. Joint probability mass function for $Y_1$ and $Y_2$: $p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2)$, $-\infty < y_1, y_2 < \infty$

properties:

1. $0 \le p(y_1, y_2) \le 1$, $\forall\, y_1, y_2$
2. $\sum_{y_1, y_2} p(y_1, y_2) = 1$

Joint bivariate distribution function $F(y_1, y_2) = P(Y_1 \le y_1, Y_2 \le y_2)$

for discrete $= \sum_{t_1=-\infty}^{y_1} \sum_{t_2=-\infty}^{y_2} p(y_1, y_2)$

for continuous $= \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f(t_1, t_2)\, dt_2\, dt_1$

## MARGINAL PROB FUNCTION

for discrete: $p_1(y_1) = \sum_{y_2} p(y_1, y_2)$, $p_2(y_2) = \sum_{y_1} p(y_1, y_2)$

for continuous:

$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2$, $f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1$

## CONDITIONAL PDF

Conditional discrete probability function: (for discrete)

$$p(y_1 \mid y_2) = P(Y_1 = y_1 \mid Y_2 = y_2)$$

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

$$= \frac{p(y_1, y_2)}{p_2(y_2)} \leftarrow \text{marginal}$$

conditional distribution function: (for continuous)

$$F(y_1 \mid y_2) = P(Y_1 \le y_1 \mid Y_2 = y_2)$$

$$= \int_{-\infty}^{y_1} \frac{f(t_1, y_2)}{f_2(y_2)} \, dt_1$$

conditional density function:

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

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

eg $f(y_1, y_2) = \begin{cases} 1/2, & 0 \le y_1 \le y_2 \le 2 \\ 0 & \text{otherwise} \end{cases}$

Find conditional density of $Y_1$ given $Y_2 = y_2$

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

$$f_2(y_2) = \int_{-\infty}^{\infty} f(y_1, y_2) \, dy_1$$

$$= \int_0^{y_2} \frac{1}{2} \, dy_1$$

$$= \frac{1}{2} y_1 \Big|_0^{y_2}$$

$$= \frac{1}{2} y_2$$

$$\Rightarrow f(y_1 \mid y_2) = \frac{\frac{1}{2}}{\frac{1}{2} y_2} = \frac{1}{y_2}, \quad 0 \le y_1 \le y_2$$

eg As 4 Q7 : $Y_1$ and $Y_2$ are continuous r.v., joint pdf:

$$f_{Y_1,Y_2}(y_1,y_2) = c\,(3y_1y_2 + y_1^2 + y_2^2), \quad 0<y_1<1,\ 0<y_2<1$$

a) Find the value of $c>0$:

idea: for $f_{Y_1,Y_2}$ to be a true pdf,

$$\int_0^1 \int_0^1 c\,(3y_1y_2 + y_1^2 + y_2^2)\,dy_2\,dy_1 = 1$$

$\Rightarrow$ solve for $c$ under that constraint:

$$c \int_0^1 \int_0^1 3y_1y_2 + y_1^2 + y_2^2\,dy_2\,dy_1$$

$$= c \int_0^1 \left[\frac{3y_1y_2^2}{2} + y_1^2y_2 + \frac{y_2^3}{3}\right]_0^1 dy_1$$

$$= c \int_0^1 \frac{3}{2}y_1 + y_1^2 + \frac{1}{3}\,dy_1$$

$$= c\left[\frac{3}{4}y_1^2 + \frac{y_1^3}{3} + \frac{1}{3}y_1\right]_0^1$$

$$= c\left[\frac{3}{4} + \frac{1}{3} + \frac{1}{3}\right]$$

$$= \frac{17}{12}c$$

$$\frac{17}{12}c = 1 \iff c = \frac{12}{17}$$

b) Find $F_{Y_1,Y_2}(y_1,y_2)$.

idea: break it down to cover all values of $y_1$, $y_2$.

1. $F_{Y_1,Y_2} = 0$ for $y_1 < 0$, $y_2 < 0$
2. $F_{Y_1,Y_2} = 1$ for $y_1 > 1$, $y_2 > 1$
3. $0<y_1<1$, $0<y_2<1$
4. $0<y_1<1$, $y_2>1$
5. $0<y_2<1$, $y_1>1$

c) Find $f_{Y_1}(y_1)$

idea: $f_{Y_1}(y_1) = \int_0^1 f_{Y_1,Y_2}\,dy_2$ for $0<y_1<1$

$f_{Y_1}(y_1) = 0$ otherwise.

# INDEPENDENCE

$Y_1, Y_2$, continuous r.v.s, are independent:

$$f_{Y_1,Y_2}(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) \quad \forall\, y_1, y_2$$

$$f_{Y_2|Y_1}(y_2 | y_1) = f_{Y_2}(y_2)$$
$$f_{Y_1|Y_2}(y_1 | y_2) = f_{Y_1}(y_1)$$

It follows too for the CDF: $F_{Y_1,Y_2}(y_1, y_2) = F_{Y_1}(y_1) \cdot F_{Y_2}(y_2)$

For discrete r.v., $p(y_1, y_2) = p(y_1) \cdot p(y_2)$.

eg $f(y_1, y_2) = \begin{cases} 6y_1 y_2^2 & , 0 \le y_1 \le 1,\ 0 \le y_2 \le 1 \\ 0 & , \text{otherwise} \end{cases}$

Show that $Y_1$ and $Y_2$ are independent.

idea: find marginal pdf of each first.

$$\begin{aligned} f(y_1) &= \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_2 \\ &= \int_0^1 6y_1 y_2^2\, dy_2 \\ &= \left[2y_1 y^3\right]_0^1 \\ &= 2y_1 \end{aligned}$$

$$f(y_1) = \begin{cases} 2y_1 & , 0 \le y_1 \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$\begin{aligned} f(y_2) &= \int_{-\infty}^{\infty} f(y_1, y_2)\, dy_1 \\ &= \int_0^1 6y_1 y_2^2\, dy_1 \\ &= \left[3y_1^2 y_2^2\right]_0^1 \\ &= 3y_2^2 \end{aligned}$$

$$f(y_2) = \begin{cases} 3y_2^2 & , 0 \le y_2 \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$\begin{aligned} f(y_1) \cdot f(y_2) &= 2y_1 \cdot 3y_2^2 \\ &= 6y_1 y_2^2 \quad \text{for } 0 \le y_1 \le 1,\ 0 \le y_2 \le 1. \end{aligned}$$

Hence, $Y_1$ and $Y_2$ are independent.

## EXPECTATIONS OF JOINT DIST

discrete: $E(g(Y_1, Y_2)) = \sum_{y_1} \sum_{y_2} g(y_1, y_2) \cdot p(y_1, y_2)$

continuous: $E(g(Y_1, Y_2)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2) \cdot f(y_1, y_2)\, dy_2\, dy_1$

If $Y_1$ and $Y_2$ are independent r.v.'s, then

$$E(g(Y_1) \cdot h(Y_2)) = E(g(Y_1)) \cdot E(h(Y_2)).$$

## COVARIANCE

$Y_1, Y_2$ are r.v.'s with means $\mu_1$ and $\mu_2$ respectively.

$$\begin{aligned} cov(Y_1, Y_2) &= E((Y_1 - \mu_1)(Y_2 - \mu_2)) \\ &= E(Y_1 Y_2) - E(Y_1)E(Y_2) \end{aligned}$$

Correlation: $corr(Y_1, Y_2) = \dfrac{cov(Y_1, Y_2)}{\sigma_1 \sigma_2}$

$$= \frac{cov(Y_1, Y_2)}{\sqrt{var(Y_1)}\sqrt{var(Y_2)}}$$

properties

1. $-1 \leq corr(Y_1, Y_2) \leq 1$
2. If $cov(Y_1, Y_2) = 0$, then $corr(Y_1, Y_2) = 0$ ie $Y_1$ and $Y_2$ are uncorrelated.

If $Y_1, Y_2$ are independent, then $cov(Y_1, Y_2) = 0$ (since $E(Y_1 Y_2) = E(Y_1)E(Y_2)$)

However note that $cov = 0 \not\Rightarrow$ independence!

## CONDITIONAL EXPECTATIONS

$Y_1, Y_2$ are 2 r.v.'s, conditional expectation of $g(Y_1)$ given $Y_2 = y_2$:

$E(g(Y_1) \mid Y_2 = y_2) = \int_{-\infty}^{\infty} g(y_1)\, f(y_1 \mid y_2)\, dy_1$ for ~~discr~~ continuous

$E(g(Y_1) \mid Y_2 = y_2) = \sum_{y_1} g(y_1)\, p(y_1 \mid y_2)$ for discrete

Law of Iterated Expectations: $E(Y_2) = E[E(Y_2 \mid Y_1)]$

## Sum of r.v.s using MGF

$m_Y(t) = E(e^{tY})$

If $U = g(Y)$ is a function on $Y$,

$m_U(t) = E(e^{tU})$

$= E(e^{t\,g(Y)})$

For 2 r.v.s $Y_1, Y_2$, where $Y = a_1 g_1(Y_1) + a_2 g_2(Y_2)$,

$m_Y(t) = E(e^{tY})$

$= E(e^{t(a_1 g_1(Y_1) + a_2 g_2(Y_2))})$

$= E(e^{t(a_1 g_1(Y_1))} \cdot e^{t(a_2 g_2(Y_2))})$

If $Y_1$ and $Y_2$ are independent, then

$E(e^{t(a_1 g_1(Y_1))} e^{t(a_2 g_2(Y_2))}) = E(e^{t(a_1 g_1(Y_1))}) E(e^{t(a_2 g_2(Y_2))})$

eg Find mgf for $Y = Y_1 + Y_2$, $Y_1 \sim \text{Poisson}(\lambda_1)$, $Y_2 \sim \text{Poisson}(\lambda_2)$, $Y_1$ and $Y_2$ are independent.

$m_{Y_1}(t) = \exp\{\lambda_1(e^t - 1)\}$

$m_{Y_2}(t) = \exp\{\lambda_2(e^t - 1)\}$

$m_Y(t) = \exp\{\lambda_1(e^t - 1)\} \cdot \exp\{\lambda_2(e^t - 1)\}$

$= \exp\{\lambda_1(e^t - 1) + \lambda_2(e^t - 1)\}$

$= \exp\{(\lambda_1 + \lambda_2)(e^t - 1)\}$

$Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$

## SAMPLE MEAN

$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$

$E(\bar{Y}_n) = \mu$ (true mean)

$Var(\bar{Y}_n) = \frac{1}{n} \sigma^2$

## CLT

$\bar{Y}$ is asymptotically normally distributed with mean $\mu$ and variance $\sigma^2/n$.
CLT can be applied for any distribution as long as $E(Y_i) = \mu$ and $Var(Y_i) = \sigma^2$ are finite, and $n$ is sufficiently large.

eg Let test scores be T, $E(T) = 60$, $Var(T) = 64$. A sample $n = 100$ of students' tests scores had a mean of 58. Are these students inferior?

Idea: calculate $P(\bar{T} \le 58)$

$$P(\bar{T} \le 58) = P\left( \frac{\bar{T} - \mu_T}{\frac{\sigma_T}{\sqrt{n}}} \le \frac{58 - 60}{\frac{\sqrt{64}}{\sqrt{100}}} \right)$$

$$= P(Z \le -2.5)$$

$$= 1 - 0.9938$$

$$= 0.0062.$$

unlikely that they are.

eg A manufacturing process produces bolts with mean diameter 0.5 inches and s.d. 0.02 inches. Each day, 36 bolts are inspected. If the resulting sample mean is $< 0.49$ inches or $> 0.51$ inches, the process stops for adjustments.

a) What is the approx dist for $\bar{Y}_{36}$?

By CLT, $\bar{Y}_n \overset{a}{\sim} N(0.5, 0.02^2)$

b) What is the P(shutdown)?

$$P(\text{shutdown}) = P(\bar{Y} < 0.49) + P(\bar{Y} > 0.51)$$

$$= P\left(Z < \frac{0.49 - 0.5}{\frac{0.02}{\sqrt{36}}}\right) + P\left(Z > \frac{0.51 - 0.5}{\frac{0.02}{\sqrt{36}}}\right)$$

$$= P(Z < -3) + P(Z > 3)$$

$$= 2P(Z < -3) \quad \text{since } Z \text{ symmetric}$$

$$= 2(0.0013)$$

$$= 0.0026$$