

## Assignment 2: CS215

Satyam Sinoliya, 23B0958  
Vaibhav Singh, 23B1068  
Shaik Awez Mehtab, 23B1080

## Solution 1

## Task A

When  $X \sim \text{Ber}(p)$ , PGF of  $X$  is

$$G_{\text{Ber}}(z) = \mathbb{E}(z^X) \quad (1)$$

$$= \sum_{n=0}^{\infty} P[X = n]z^n \quad (2)$$

Since  $P[X = 0] = (1 - p)$ ,  $P[X = 1] = p$ ,  $P[X = n] = 0$  when  $n > 1$ ,

$$G_{\text{Ber}}(z) = P[X = 0]z^0 + P[X = 1]z^1 \quad (3)$$

$$= (1 - p) + pz \quad (4)$$

## Task B

When  $X \sim \text{Bin}(n, p)$ , PMF of  $X$  is

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k} \text{ for } k \leq n. \quad (5)$$

and  $P[X = k] = 0$  for  $k > n$ .

$$G_{\text{Bin}}(z) = \sum_{k=0}^{\infty} P[X = k]z^k \quad (6)$$

$$= \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} z^k \quad (7)$$

$$= \sum_{k=0}^n \binom{n}{k} (pz)^k (1 - p)^{n-k} \quad (8)$$

$$= (1 - p + pz)^n. \quad (9)$$

By equation 4,  $G_{\text{Bin}}(z) = (1 - p + pz)^n = (G_{\text{Ber}}(z))^n$ . Hence proved.

## Task C

By the definition of PGF,

$$G(z) = \sum_{n=0}^{\infty} P[X_1 = n]z^n \quad (10)$$

Let  $(G(z))^k = \sum_{n=0}^{\infty} a_n z^n$ . Now,  $G_{\Sigma}(z)$  is

$$G_{\Sigma}(z) = \sum_{n=0}^{\infty} P[X = n]z^n \quad (11)$$

$$= \sum_{n=0}^{\infty} P[X_1 + \dots + X_k = n]z^n \quad (12)$$

$$= \sum_{n=0}^{\infty} \sum P[X_1 = i_1, \dots, X_k = i_k]z^n \quad (13)$$

where  $i_1 + \dots + i_k = n$ .

$$= \sum_{n=0}^{\infty} \sum P(i_1) \dots P(i_n) z^n \quad (14)$$

Now,  $G(z) = \sum_{n=0}^{\infty} P(n)z^n$ . And since  $a_n$  is coefficient of  $z^n$  in  $(G(z))^k = (\sum_{n=0}^{\infty} P(n)z^n)^k$ .

$$a_n = \sum P(i_1)P(i_2) \dots P(i_k) \text{ where } i_1 + \dots + i_k = n \quad (15)$$

By equation 14

$$G_{\Sigma}(z) = \sum_{n=0}^{\infty} a_n z^n = (G(z))^k \quad (16)$$

Hence Proved.

#### Task D

When  $X \sim \text{Geo}(p)$ , PMF of  $X$ ,

$$P[X = k] = (1 - p)^{k-1}p \quad (17)$$

for  $k > 0$ .  $P[X = 0] = 0$ . Now, PGF of  $X$ ,

$$G_{\text{Geo}}(z) = \sum_{k=0}^{\infty} P[X = k]z^k \quad (18)$$

$$= \sum_{k=1}^{\infty} P[X = k]z^k \quad (19)$$

$$= \sum_{k=1}^{\infty} p(1 - p)^{k-1}z^k \quad (20)$$

$$= \sum_{k=1}^{\infty} pz(z - zp)^{k-1} \quad (21)$$

$$= pz \sum_{k=0}^{\infty} (z - zp)^k \quad (22)$$

$$= \frac{pz}{1 - z + pz} \quad (23)$$

#### Task E

By equation 9,

$$G_{\text{Bin}}(z) = (1 - p + pz)^n = G_X^{(n,p)}(z). \quad (24)$$

For  $Y \sim \text{NegBin}(n, p)$ ,  $Y$  represents the number of independent coin throws required to get  $n$  heads of a coin. Let  $X_i$  represents the number of throws of coin required after getting  $(i - 1)^{\text{th}}$  head to get the  $i^{\text{th}}$  head. Since all of the coin throws are independent, the outcome of a given throw doesn't depend on the previous coins' output. Thus,  $X_i$  is just the number of throws to get a head when a coin is thrown, where each  $X_i \sim \text{Geo}(p)$  since each coin is same with probability of getting head as  $p$ .

$Y$  can be written as  $Y = X_1 + X_2 + \dots + X_k$ . Using equations 16 and 23,

$$G_Y^{(n,p)}(z) = (G_{\text{Geo}}(z))^n \quad (25)$$

$$= \left( \frac{pz}{1 - z + pz} \right)^n \quad (26)$$

$$(27)$$

$$G_X^{(n,p^{-1})}(z^{-1}) = \left(1 - \frac{1}{p} + \frac{1}{pz}\right)^n \quad (28)$$

$$= \left(\frac{1 - z + pz}{pz}\right)^n \quad (29)$$

$$\left(G_X^{(n,p^{-1})}(z^{-1})\right)^{-1} = \left(\frac{pz}{1 - z + pz}\right)^n \quad (30)$$

$$= G_Y^{(n,p)}(z) \quad (31)$$

Hence Proved.

#### Task F

For  $Y \sim \text{NegBin}(n, p)$ ,

$$P[Y = k] = \binom{k-1}{n-1} p^n (1-p)^{k-n} \text{ for } k \geq n \quad (32)$$

Otherwise,  $P[Y = k] = 0$ . PGF of  $Y$  is

$$G_Y^{(n,p)}(z) = \sum_{k=0}^{\infty} P[Y = k] z^k \quad (33)$$

$$= \sum_{k=n}^{\infty} \binom{k-1}{n-1} p^n (1-p)^{k-n} z^k \quad (34)$$

$$= \sum_{k=0}^{\infty} \binom{k+n-1}{n-1} p^n (1-p)^k z^{n+k} \quad (35)$$

$$= (pz)^n \sum_{k=0}^{\infty} \binom{k+n-1}{n-1} (z-pz)^k \quad (36)$$

Using equation 27,

$$\left(\frac{pz}{1-z+pz}\right)^n = (pz)^n \sum_{k=0}^{\infty} \binom{k+n-1}{n-1} (z-pz)^k \quad (37)$$

$$(1 - (z-pz))^{-n} = \sum_{k=0}^{\infty} \binom{k+n-1}{n-1} (z-pz)^k \quad (38)$$

Since  $z, p$  are arbitrary, let  $z - pz = x$ .

$$(1-x)^{-n} = \sum_{r=0}^{\infty} \binom{r+n-1}{n-1} x^r = \sum_{r=0}^{\infty} \binom{-n}{r} x^r \quad (39)$$

Hence proved.

#### Task G

**To prove:** Given PGF of a random variable  $X$  is  $G(z)$ , expectation of  $X$  i.e  $\mathbb{E}(x) = G'(1)$

**Proof:**

$$G(z) = \mathbb{E}(z^X) = \sum_{k=0}^{\infty} P[X = k] z^k \quad (40)$$

$$G'(z) = \sum_{k=0}^{\infty} k P[X = k] z^{k-1} \quad (41)$$

$$G'(1) = \sum_{k=0}^{\infty} k P[X = k] \quad (42)$$

$$= \mathbb{E}[X] \quad (43)$$

Hence Proved. Now, Let's derive means of Bernoulli, Binomial, Geometric and Negative Binomial distributions using this:

1. **Bernoulli Distribution:** Let  $X \sim \text{Ber}(p)$ ,

$$G_{\text{Ber}}(z) = (1 - p) + pz \quad (44)$$

$$G'_{\text{Ber}}(z) = p \quad (45)$$

$$G'_{\text{Ber}}(1) = p = \mathbb{E}[X] \quad (46)$$

Thus,  $\mathbb{E}[X] = p$ .

2. **Binomial Distribution:** Let  $X \sim \text{Bin}(n, p)$ ,

$$G_{\text{Bin}}(z) = (1 - p + pz)^n \quad (47)$$

$$G'_{\text{Bin}}(z) = np(1 - p + pz)^{n-1} \quad (48)$$

$$G'_{\text{Bin}}(1) = np = \mathbb{E}[X] \quad (49)$$

Thus,  $\mathbb{E}[X] = np$ .

3. **Geometric Distribution:** Let  $X \sim \text{Geo}(p)$ ,

$$G_{\text{Geo}}(z) = \frac{pz}{1 - z + pz} \quad (50)$$

$$G'_{\text{Geo}}(z) = \frac{p(1 - z + pz) - pz(p - 1)}{(1 - z + pz)^2} \quad (51)$$

$$= \frac{p}{(1 - z + pz)^2} \quad (52)$$

$$G'_{\text{Geo}}(1) = \frac{p}{p^2} = \frac{1}{p} = \mathbb{E}[X] \quad (53)$$

Thus,  $\mathbb{E}[X] = \frac{1}{p}$ .

4. **Negative Binomial Distribution:** Let  $X \sim \text{NegBin}(n, p)$ ,

$$G_{\text{NegBin}}(z) = \left( \frac{pz}{1 - z + pz} \right)^n \quad (54)$$

$$G'_{\text{NegBin}}(z) = n \left( \frac{pz}{1 - z + pz} \right)^{n-1} \left( \frac{p}{(1 - z + pz)^2} \right) \quad (55)$$

$$G'_{\text{NegBin}}(1) = \frac{n}{p} = \mathbb{E}[X] \quad (56)$$

Thus,  $\mathbb{E}[X] = \frac{n}{p}$ .

## Solution 2

### Task A

#### To prove:

Let  $X$  be a continuous real-valued random variable with CDF  $F_X : \mathbb{R} \rightarrow [0, 1]$ . Assume that  $F_X$  is invertible. Then the random variable  $Y := F_X(X) \in [0, 1]$  is uniformly distributed in  $[0, 1]$

#### Proof:

$F_X$  by definition can also be written as

$$F_X(x) = P(X \leq x)$$

Define a new random variable  $Y$ ,

$$Y = F_X(X)$$

$Y$  is the result of applying CDF  $F_X$  to the random variable  $X$ . To prove the theorem, assume  $y \in [0, 1]$ . So, the probability that  $Y \leq y$  is:

$$P(Y \leq y) = P(F_X(X) \leq y)$$

It is assumed that  $F_X(x)$  is invertible, so,

$$P(Y \leq y) = P(X \leq F_X^{-1}(y))$$

which is basically, probability that  $X$  is less than or equal to  $F_X^{-1}(y)$ . This can be written in the CDF form, which is  $F_X(F_X^{-1}(y))$ . So,

$$P(Y \leq y) = P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y$$

So,

$$P(Y \leq y) = y$$

where  $y \in [0, 1]$ , which is the CDF of uniform distribution in  $[0, 1]$ . So,  $Y$  is a uniform distribution in  $[0, 1]$  regardless of  $X$ .

#### Task B

According to the theorem proved above, CDF of any random variable  $X$  mapped with itself gives a uniform random variable  $Y$  in  $[0, 1]$ . So, let  $Y \sim \text{Uniform}(0, 1)$ . Then for any random variable  $X$ ,

$$F_X(X) = Y \quad (57)$$

$$X = F_X^{-1}(Y) \quad (58)$$

#### Algorithm $\mathcal{A}$ :

1. Input: A sample  $y$  from the uniform distribution on  $[0, 1]$ .
2. Transformation:
  - Apply the inverse CDF to  $y$  to compute a sample  $u$ .
  - Define  $\mathcal{A}(u) = u = F_X^{-1}(y)$
3. Output: The random variable  $U = F_X^{-1}(Y)$

This gives us the correct required random variables as, CDF of  $U$  is  $F_U(u)$ ,

$$P(U \leq u) = P(F_X(Y) \leq u) \quad (59)$$

$$F_U(u) = P(F_X(F_X^{-1}(X)) \leq u) \quad (60)$$

$$F_U(u) = P(X \leq u) \quad (61)$$

$$F_U(u) = F_X(u) \quad (62)$$

$$(63)$$

$U$  and  $X$  have the same CDF, which was initially required.

## Solution 3

## Task D

Given, PDF of Gamma-distribution  $\text{Gamma}(k, \theta)$  is  $f(x; k, \theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}}$ . First moment of it is

$$\mu_1^{\text{Gamma}} = \mathbb{E}[X] \quad (64)$$

$$= \int_{-\infty}^{\infty} x \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}} dx \quad (65)$$

$$= \frac{1}{\theta^k \Gamma(k)} \int_{-\infty}^{\infty} x^k e^{-\frac{x}{\theta}} dx \quad (66)$$

Let  $u = \frac{x}{\theta}$ , then  $\theta du = dx$

$$= \frac{\theta}{\Gamma(k)} \int_{-\infty}^{\infty} u^k e^{-u} du \quad (67)$$

$$= \frac{\theta \Gamma(k+1)}{\Gamma(k)} \quad (68)$$

Since  $\Gamma(k+1) = k\Gamma(k)$

$$= k\theta \quad (69)$$

Second moment of it is

$$\mu_2^{\text{Gamma}} = \mathbb{E}[X^2] \quad (70)$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}} dx \quad (71)$$

$$= \frac{1}{\theta^k \Gamma(k)} \int_{-\infty}^{\infty} x^{k+1} e^{-\frac{x}{\theta}} dx \quad (72)$$

Let  $u = \frac{x}{\theta}$ , then  $\theta du = dx$

$$= \frac{\theta^2}{\Gamma(k)} \int_{-\infty}^{\infty} u^{k+1} e^{-u} du \quad (73)$$

$$= \frac{\theta^2 \Gamma(k+2)}{\Gamma(k)} \quad (74)$$

$$= (k+1)k\theta^2 \quad (75)$$

Thus,  $\mu_1^{\text{Gamma}} = k\theta$ ,  $\mu_2^{\text{Gamma}} = (k+1)k\theta^2$ .

## Solution 4

## Solution 5

## Task A

Given, PDF of GMM variable  $X$  is  $f_X = \sum_{i=1}^K p_i P[X_i = x]$ . Let it's CDF be  $F_X$ . Then  $F_X(x)$  is given by

$$F_X(x) = P[X \leq x] \quad (76)$$

$$= \int_{-\infty}^x f_X(t) dt \quad (77)$$

$$= \int_{-\infty}^x \sum_{i=1}^K p_i P[X_i = t] dt \quad (78)$$

$$= \sum_{i=1}^K p_i \int_{-\infty}^x P[X_i = t] dt \quad (79)$$

$$= \sum_{i=1}^K p_i P[X_i \leq x] \quad (80)$$

$$= \sum_{i=1}^K p_i F_{X_i}(x) \quad (81)$$

Where  $F_{X_i}(x) = P[X_i \leq x]$  is CDF of  $X_i$ .

Now, let CDF of output of the given algorithm be  $F_A(x) = P[\mathcal{A} \leq x]$ . Since the events that we choose  $\mathcal{A}$  to be from the distribution  $i$  (say  $E_i$ ) are disjoint for  $i = 1, \dots, k$ .

$$F_A(x) = P[\mathcal{A} \leq x] \quad (82)$$

$$= \sum_{i=1}^K P[E_i] \cdot P[\mathcal{A} \leq x | E_i] \quad (83)$$

$$= \sum_{i=1}^K p_i F_{X_i}(x) \quad (84)$$

$$= F_X(x) \quad (85)$$

We know that PDF of a random variable  $X$  with CDF  $F_X(x)$  is  $\frac{\partial F_X}{\partial x}$ . Thus,

$$f_A(x) = \frac{\partial F_A}{\partial x} \quad (86)$$

$$= \frac{\partial F_X}{\partial x} \quad (87)$$

$$= f_X \quad (88)$$

Since  $x$  was arbitrary, for every  $u \in \mathbb{R}$ ,  $f_A(u) = f_X(u)$ . i.e the algorithm indeed samples from the GMM variable's distribution.

### Task B

Since

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} t \cdot P[X = t] dt \quad (89)$$

$$= \int_{-\infty}^{\infty} t \cdot \sum_{i=1}^K p_i P[X_i = t] dt \quad (90)$$

$$= \sum_{i=1}^K p_i \int_{-\infty}^{\infty} t P[X_i = t] dt \quad (91)$$

$$= \sum_{i=1}^K p_i \mathbb{E}[X_i] \quad (92)$$

$$= \sum_{i=1}^K p_i \mu_i \quad (93)$$



Let  $\mu = \mathbb{E}[X]$ .

$$\text{Var}[X] = \int_{-\infty}^{\infty} (t - \mu)^2 P[X = t] dt \quad (94)$$

$$= \int_{-\infty}^{\infty} (t - \mu)^2 \sum_{i=1}^K p_i P[X_i = t] dt \quad (95)$$

$$= \sum_{i=1}^K p_i \int_{-\infty}^{\infty} (t - \mu)^2 P[X_i = t] dt \quad (96)$$

$$= \sum_{i=1}^K p_i \text{Var}[X_i] \quad (97)$$

$$= \sum_{i=1}^K p_i \sigma_i^2 \quad (98)$$

Let  $\sigma^2 = \text{Var}[X]$ .

$$\text{MGF}_X(t) = \int_{-\infty}^{\infty} e^{tX} P[X = x] dx \quad (99)$$

$$= \int_{-\infty}^{\infty} e^{tX} \sum_{i=1}^K p_i P[X_i = x] dx \quad (100)$$

$$= \sum_{i=1}^K p_i \int_{-\infty}^{\infty} e^{tX} P[X_i = x] dx \quad (101)$$

$$= \sum_{i=1}^K p_i \text{MGF}_{X_i}(t) \quad (102)$$

$$= \sum_{i=1}^K p_i e^{t\mu_i + \frac{1}{2}t^2\sigma_i^2} \quad (103)$$

### Task C

Given  $Z = \sum_{i=1}^K p_i X_i$ , where  $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{i=1}^K p_i X_i\right] \quad (104)$$

$$= \sum_{i=1}^K p_i \mathbb{E}[X_i] \quad (105)$$

$$= \sum_{i=1}^K p_i \mu_i \quad (106)$$