CS215 Assignment 2

Kavya Gupta

September 2023

Solution for Q4

Proving $P(X \geq x) \leq e^{-tx}\Phi_X(t), t > 0$

When X is a Discrete Random Variable

For a discrete r.v., its MGF

$$\Phi_X(t) = E(e^{tX}) = \sum_{i=1}^{n} e^{tx_i} P(X = x_i) \tag{1}$$

Where x_is are the outcomes of X. So any real x lies in $(-\infty, x_1]$ or (x_n, ∞) or $(x_i, x_{i+1}]$ for some i.
If $x \in (x_n, \infty)$, then there exists no x_i such that $X \geq x$, so trivially $P(X \geq x) = 0$ which is definitely lesser than RHS as RHS ≥ 0. Hence this case is trivially proved.
If $x \in (x_i, x_{i+1}]$, then $x_k \geq x$ for all $k \in \{i+1, i+2, ..., n\}$, and for $x \in (-\infty, x_1]$, $x_k \geq x$ for all $k \in \{1, 2, ..., n\}$, hence :-

$$P(X \geq x) = \sum_{r=i+1}^{n} P(X = x_r)$$

Also, $e^{t(x_r - x)} \geq 1$ as $t > 0$ and $x_r \geq x$, hence :-

$$P(X \geq x) = \sum_{r=i+1}^{n} P(X = x_r) \leq \sum_{r=i+1}^{n} e^{t(x_r - x)} P(X = x_r) = e^{-tx} \sum_{r=i+1}^{n} e^{tx_r} P(X = x_r)$$

It will be less than or equal to summation from 1 to n as all the terms are non-negative.

$$\Rightarrow P(X \geq x) \leq e^{-tx} \sum_{r=i+1}^{n} e^{tx_r} P(X = x_r) \leq e^{-tx} \sum_{r=1}^{n} e^{tx_r} P(X = x_r)$$

Comparing with MGF, we see that :-

$$P(X \geq x) \leq e^{-tx} \sum_{r=1}^{n} e^{tx_r} P(X = x_r) = e^{-tx}\Phi_X(t) \Rightarrow P(X \geq x) \leq e^{-tx}\Phi_X(t)$$

Hence proved for discrete r.v.

When X is a Continuous Random Variable

MGF of a continuous r.v. is

$$\Phi_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx \tag{2}$$

Where $f_X(x)$ is the pdf of X, also :-

$$P(X \geq x_0) = \int_{x_0}^{\infty} f_X(x)\, dx$$

Again $e^{t(x-x_0)} \geq 1$ as $t > 0$ and $x \geq x_0$ in the above integral, so $f_X(x) \leq e^{t(x-x_0)} f_X(x)$ for all $x \geq x_0$, hence :-

$$P(X \geq x_0) = \int_{x_0}^{\infty} f_X(x)\, dx \leq \int_{x_0}^{\infty} e^{t(x-x_0)} f_X(x)\, dx = e^{-tx_0} \int_{x_0}^{\infty} e^{tx} f_X(x)\, dx$$

Also, since $e^{tx} f_X(x) \geq 0$ for all x, the above will be $\leq$ integral from $-\infty$ to ∞.

$$P(X \geq x_0) \leq e^{-tx_0} \int_{x_0}^{\infty} e^{tx} f_X(x)\, dx \leq e^{-tx_0} \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx = e^{-tx_0} \Phi_X(t)$$

Hence this is proved for continuous random variable as well, for $t > 0$:-

$$P(X \geq x) \leq e^{-tx} \Phi_X(t)$$

Proving $P(X \leq x) \leq e^{-tx} \Phi_X(t), t < 0$

When X is a Discrete Random Variable

So any real x lies in $(-\infty, x_1)$ or $[x_n, \infty)$ or $[x_i, x_{i+1})$ for some i.
If $x \in (-\infty, x_1)$, then there exists no x_i such that $X \leq x$, so trivially $P(X \leq x) = 0$ which is definitely lesser than RHS as RHS ≥ 0. Hence this case is trivially proved.
If $x \in [x_i, x_{i+1})$, then $x_k \leq x$ for all $k \in \{1, 2, ..., i\}$, and for $x \in [x_n, \infty)$, $x_k \leq x$ for all $k \in \{1, 2, ..., n\}$, hence :-

$$P(X \leq x) = \sum_{r=1}^{i} P(X = x_r)$$

Also, $e^{t(x_r - x)} \geq 1$ as $t < 0$ and $x_r \leq x$, hence :-

$$P(X \leq x) = \sum_{r=1}^{i} P(X = x_r) \leq \sum_{r=1}^{i} e^{t(x_r - x)} P(X = x_r) = e^{-tx} \sum_{r=1}^{i} e^{tx_r} P(X = x_r)$$

It will be less than or equal to summation from 1 to n as all the terms are non-negative.

$$\Rightarrow P(X \leq x) \leq e^{-tx} \sum_{r=1}^{i} e^{tx_r} P(X = x_r) \leq e^{-tx} \sum_{r=1}^{n} e^{tx_r} P(X = x_r)$$

Comparing with MGF (from (1)), we see that :-

$$P(X \leq x) \leq e^{-tx} \sum_{r=1}^{n} e^{tx_r} P(X = x_r) = e^{-tx}\Phi_X(t) \Rightarrow P(X \leq x) \leq e^{-tx}\Phi_X(t)$$

Hence proved for discrete r.v.

When X is a Continuous Random Variable

Considering $f_X(x)$ is the pdf of X, then :-

$$P(X \leq x_0) = \int_{-\infty}^{x_0} f_X(x)\, dx$$

Again $e^{t(x-x_0)} \geq 1$ as $t < 0$ and $x \leq x_0$ in the above integral, so $f_X(x) \leq e^{t(x-x_0)} f_X(x)$ for all $x \leq x_0$, hence :-

$$P(X \leq x_0) = \int_{-\infty}^{x_0} f_X(x)\, dx \leq \int_{-\infty}^{x_0} e^{t(x-x_0)} f_X(x)\, dx = e^{-tx_0} \int_{-\infty}^{x_0} e^{tx} f_X(x)\, dx$$

Also, since $e^{tx} f_X(x) \geq 0$ for all x, the above will be $\leq$ integral from $-\infty$ to ∞, comparing MGF from (2).

$$P(X \leq x_0) \leq e^{-tx_0} \int_{-\infty}^{x_0} e^{tx} f_X(x)\, dx \leq e^{-tx_0} \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx = e^{-tx_0}\Phi_X(t)$$

Hence this is proved for continuous random variable as well, for $t < 0$:-

$$P(X \leq x) \leq e^{-tx}\Phi_X(t)$$

Now other part of the question...

Given that X is a sum of n independent Bernoulli variables $X_1, X_2, ..., X_n$. Each Bernoulli variable is a discrete r.v, hence its sum = X is also a discrete r.v.
Let's use $P(X \geq x) \leq e^{-tx}\Phi_X(t), t \geq 0$ proved above for discrete r.v.
Put $(1+\delta)\mu$ as x in this formula, we get:-

$$P(X \geq (1+\delta)\mu) \leq e^{-t(1+\delta)\mu}\Phi_X(t) \tag{3}$$

We know that for a Bernoulli variable X_i, its MGF is (where $E[X_i] = p_i$):-

$$\Phi_{X_i}(t) = 1 - p_i + p_i e^t$$

Also, as $X = \sum_{i=1}^{n} X_i$ and all X_is are independent, then:-

$$\Phi_X(t) = \prod_{i=1}^{n} \Phi_{X_i}(t) = \prod_{i=1}^{n} (1 - p_i + p_i e^t)$$

Here we used that if $X = X_1 + X_2$ and X_1, X_2 are independent, then $\Phi_X(t) = \Phi_{X_1}(t).\Phi_{X_2}(t)$
Using $1 + x \leq e^x$, put $x = p_i(e^t - 1)$:-

$$1 - p_i + p_i e^t = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}$$

$$\Rightarrow \Phi_X(t) = \prod_{i=1}^{n}(1 - p_i + p_i e^t) \leq \prod_{i=1}^{n} e^{p_i(e^t-1)} = e^{\sum_{i=1}^{n} p_i(e^x-1)}$$

$$\Rightarrow \Phi_X(t) = e^{(e^t-1)\sum_{i=1}^{n} p_i} = e^{\mu(e^t-1)}$$

Put this in equation (3)...

$$P(X \geq (1+\delta)\mu) \leq e^{-t(1+\delta)\mu}\Phi_X(t) \leq e^{-t(1+\delta)\mu}e^{\mu(e^t-1)} = \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)t\mu}}$$

Final Touch...

$$P(X > (1+\delta)\mu) \leq P(X \geq (1+\delta)\mu) \Rightarrow P(X > (1+\delta)\mu) \leq \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)t\mu}}$$

Hence formula proved.

Tightening the bound...

Since the above formula is true for all $t \geq 0$, it should be true for that t_0 too for which RHS is minimum. Hence we minimize the RHS:-

$$RHS = g(t) = \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)t\mu}} = e^{\mu(e^t-1)-(1+\delta)t\mu} = e^{\mu(e^t-(1+\delta)t-1)}$$

Since e^x is an increasing function, minimum value of expression exists at minimum value of power. Differentiating $f(t) = e^t - (1+\delta)t - 1$:-

$$f'(t) = \mu(e^t - (1+\delta)), f''(t) = e^t > 0$$

$$f'(t_0) = 0 \Rightarrow e^{t_0} - (1+\delta) \Rightarrow t_0 = \ln(1+\delta)$$

Hence $f(t_0)$ is minimum or $g(t_0)$ is also minimum as stated above...

$$g(t_0) = g(\ln(1+\delta)) = e^{\mu(1+\delta-(1+\delta)\ln(1+\delta)-1)} = \frac{e^{\mu\delta}}{(1+\delta)^{\mu(1+\delta)}}$$

Finally...

$$P(X > (1+\delta)\mu) \leq \frac{e^{\mu\delta}}{(1+\delta)^{\mu(1+\delta)}}$$