1 问题一

1、记号：第一次订单报数为 P，总的订单报数为 G，收益为 y，图书定价为 m，折扣为 n，一次印刷的成本为 s，一次印刷的印刷数为 t，印刷的印张为 r.

2、假设：(1) 每年的第一次订单报数 $P(n)$ 和每年的总的订单报数 $G(n)$ 与上一年的第一次订单报数 $P(n-1)$ 及上一年的总的订单报数为 $G(n-1)$ 之间的差分具有以下的线性关系：

$$P(n) - G(n-1) = k(G(n-1) - P(n-1)) + d. \tag{1}$$

$$G(n) - P(n) = k(P(n) - G(n-1)) + d. \tag{2}$$

(2)A 类图书被退货的情况可以忽略不计.

3、用时间序列分析得到 2021 秋第一次订单报数为 $p = P(2021)$，以及 2021 秋总的订单报数为 $g = G(2021)$，它们是两个随机变量，其随机性体现在式 (1)(2) 中的系数 d 被随机变量 ε 取代，也即：

$$p - G(2020) = k(G(2020) - P(2020)) + \varepsilon. \tag{3}$$

$$g - p = k(p - G(2020)) + \varepsilon. \tag{4}$$

其中，根据假设 (1) 再利用往年的 P 与 G 的数据可以采用最小二乘法得到参数 k 和随机变量 ε 的分布，ε 的概率密度函数记为 $f(\varepsilon)$，那么 2021 秋第一次订单报数和总的订单报数的分布都可以得到，记 2021 秋第一次订单报数 p 的概率密度函数为 $f(p)$.

要求最好将印刷次数控制在 3 次以内，显然根据往年的情况，只要每次印刷的册数合适，一般来说 3 次以内的印刷是可以满足需求的. 于是这里不妨假设 2021 秋的 3 次印刷指标为 (a_1, a_2, a_3)（3 个在一定条件下变动的参数，可以取值为 0），每次印刷的册数为 (a_1p, a_2p, a_3p)，对应印制成本为 (s_1, s_2, s_3)，于是问题的关键就在于确定 (a_1, a_2, a_3).

4、根据所给的印刷成本计算表格，提取数据拟合得到一个成本计算公式为：

$$s = \begin{cases} 0.32rt, & t \geq 10000, \text{单色} \\ 0.34rt, & t < 10000, \text{单色} \\ 0.35rt, & t \geq 10000, \text{双色} \\ 0.41rt, & t < 10000, \text{双色} \\ 0.45rt, & t \geq 10000, \text{四色} \\ 0.54rt, & t < 10000, \text{四色} \end{cases} \tag{5}$$

则根据 (5) 式，得到 2021 年秋 A 类书籍的三次印制的成本分别为：

$$s_i = \begin{cases} 0.35a_ipr, & a_ip \geq 10000 \\ 0.41a_ipr, & a_ip < 10000 \end{cases} \tag{6}$$

其中，$i = 1, 2, 3$.

5、确定参数 (a_1, a_2, a_3) 的取值范围 Γ（怎样的一个范围呢？至少要求 $(a_1 + a_2 + a_3)p \geq g$），让其在其中变动. 对于任意给定的一组 $\gamma = (a_1, a_2, a_3) \in \Gamma$ 的值，当得到 2021 秋第一次订单报数 p 后，我们的印刷计划就能确定，这样就有一个相应的收益 $y_\gamma(p), \gamma \in \Gamma$. 而收益的计算公式为：

$$y_\gamma(p) = mng - (s_1 + s_2 + s_3) - 0.0273mg. \tag{7}$$

6、由于 p 未知，为了减少风险，我们在平均的意义下考虑收益，即针对所有可能的 p 取收益的期望作为一组给定的参数 $\gamma = (a_1, a_2, a_3)$ 的收益:

$$y_\gamma = E(y_\gamma(p)) = \int_{-\infty}^{+\infty} y_\gamma(p) f(p) \mathrm{d}x. \tag{8}$$

7、若对于所有的 $\gamma = (a_1, a_2, a_3) \in \Gamma$，$y_\gamma$ 的上确界存在，则此时的 $\gamma = (a_1, a_2, a_3)$ 能使得在本模型下收益最大，可作为 A 类图书印刷的策略，即第一次印刷 $a_1 p$ 册，第二次印刷 $a_2 p$ 册，第三次印刷 $a_3 p$ 册.

2 问题二

对于 B 类图书，由于更新比较快，当年不能销售的书将成为滞销库存，等待报废. 因此，我们在考虑 B 类图书的印刷次数、时间和数量时，不能和 A 类图书一样只考虑收益，还应该考虑库存带来的成本. 同时，由于 B 类图书的销量不像 A 类图书那样稳定，因此我们也不能事先决定在哪些月份印刷和印刷多少次，而只能在图书的畅销月份中根据销量和已有库存的数量关系去选择该月是否印刷和决定印刷量. 我们假设 2021 年销售的 B 类图书并未换新版，还是印刷已知的最新版本；且由于非最新版次的图书都已经报废，我们假设它们的销量可以忽略不计.

已知 B1 是高考复习第一轮用书，一个销售周期基本是当年 3 月到当年 9 月，而 B2 是高考复习的第二轮用书，一个销售周期基本是当年 10 月到次年 4 月，其印刷周期是跨年的，二者的销售周期都为 7 个月. 为了确定印刷的次数，我们设印刷次数为 a，在实际操作中将其设置为一个随机数，取值范围为 1-7 之间的一个整数. 设印刷的具体月份分别为 $b_1, b_2, ..., b_a$，那么根据畅销月份的不同，对于图书 B1，$b_i, i = 1, 2, ..., a$ 的取值分别为 3-9 之间的 a 个互异整数；而对于图书 B2，$b_i, i = 1, 2, ..., a$ 的取值分别为 10-12 以及 1-4（这里的 1-4 是指下一年的 1-4 月）之间的 a 个互异整数，自然地，未被选中的月份的印刷量为 0.

我们先通过时间序列分析得到 2021 年 B 类图书在每个月的销量 $p_1, p_2, ..., p_{12}$ 为 12 个随机变量，及其概率密度函数 $f_1, f_2, ..., f_{12}$，服从正态分布，记其均值分别为 $\mu_1, \mu_2, ..., \mu_{12}$，标准差分别为 $\sigma_1, \sigma_2, ..., \sigma_{12}$，我们假设 2021 年每个月的销量是近似独立的.

设最新版图书在 2021 年以前的总销量为 k，在 2021 年以前的总印刷量为 l，且在 2021 年每月的印刷量记为 $h_1, h_2, ..., h_{12}$. 我们预先设置一个常数 d 和 a 个比例系数 $c_1, c_2, ..., c_a$，对某个固定的月份 $b_N, N = 1, 2, ..., a$，若该月前的总印刷量

$$h = l + h_1 + h_2 + \cdots + h_{b_{N-1}},$$

与总销量

$$p = k + p_1 + p_2 + \cdots + p_{b_N},$$

之间的差大于 d，则表示该月无需印刷，即 $h_{b_N}=0$，若该月前的总印刷量 h 与总销量 p 之间的差小于等于 d，则认为该月需要印刷，且印刷量为:

$$h_{b_N}=c_N(h-p).$$

这样，2021 年每次印刷的印刷量我们也表示了出来.

为了确定每次印刷最优的印刷数量，我们需要先构造一个目标函数来决定怎样的印刷策略才是最优的. 对于 B 类图书，目标函数应该包含利润和库存成本两个指标. 对于利润，我们只需要考虑整个 2021 年的总利润 y，同第一题，设图书的定价为 m，折扣为 n，一次印刷的成本为 s，印刷的印张为 r. 则对 B 类图书在 b_N 月的印刷成本为：

$$s_{b_N}=\begin{cases}0.32rh_{b_N}, & h_{b_N}\geq 10000\\ 0.34rh_{b_N}, & h_{b_N}<10000\end{cases} \tag{9}$$

因而 2021 年总的利润的公式为：

$$\begin{aligned} y(p_1+p_2+\cdots+p_{12})=&mn(p_1+p_2+\cdots+p_{12})\\ &-(s_{b_1}+s_{b_2}+\cdots+s_{b_a})\\ &-0.0273m(p_1+p_2+\cdots+p_{12}).\end{aligned} \tag{10}$$

同时，根据假设，随机变量 $p^*=p_1+p_2+\cdots+p_{12}$ 服从正态分布 $N(\mu_1+\mu_2+\cdots+\mu_{12},\sigma_1^2+\sigma_2^2+\cdots+\sigma_{12}^2)$，记这个正态分布的概率密度函数为 $f(p^*)$.

这样，我们就能类似于问题一的处理方式，从平均的意义上来考虑利润，其公式为：

$$y=E(y(p^*))=\int_{-\infty}^{+\infty}y(p^*)f(p^*)\mathrm{d}x. \tag{11}$$

另一方面，为减小库存，我们应该将库存积累带来的成本合理地在目标函数中表达. 设 2021 年每月的库存量为 $w_1,w_2,...,w_{12}$，则

$$w_i(p_1+p_2+\cdots+p_i)=l-k+(h_1+h_2+\cdots+h_{i-1})-(p_1+p_2+\cdots+p_i),$$

其中，$i=1,2,...,12$. 我们定义每月的库存成本 z_i 如下：

$$z_i=0.05mw_i. \tag{12}$$

同样，w_i 是一个随机变量，我们依旧需要在平均的意义下去考虑它，由于随机变量 $p_1+p_2+\cdots+p_i$ 服从正态分布 $N(\mu_1+\mu_2+\cdots+\mu_i,\sigma_1^2+\sigma_2^2+\cdots+\sigma_i^2)$，记这个正态分布为 $f_i^*(p_1+p_2+\cdots+p_i)$，于是我们得到可量化的每月库存成本：

$$w_i=E(w_i(p_1+p_2+\cdots+p_i))=\int_{-\infty}^{+\infty}f_i^*(p_1+p_2+\cdots+p_i)z_i\mathrm{d}x, i=1,2,...,12. \tag{13}$$

我们定义目标函数 J 为利润与 12 个月的库存成本之差，即：

$$J=y-\sum_1^{12}w_i. \tag{14}$$

于是，我们的目标就是寻找使得目标函数 (14) 最大化的整数 a，以及 a 个月份数 $b_1,b_2,...,b_a$ 和 a 个比例系数 $c_1,c_2,...c_a$.

3 问题三

由于 C 类图书定位于长销，销售时间持续很多年，所以为了根据图书的热度确定在什么情况下进行下一次印刷和印刷量是多少，我们不妨取未来一年的销量来进行热度分析. 类似于问题二的记号，设 C 类图书在 2021 年 3 月 31 日以前的总销量为 k，在 2021 年 3 月 31 日以前的总印刷量为 l，印刷的印张为 r. 通过时间序列分析我们预测出 2021 年 4 月至 2022 年 3 月这 12 个月的销量，记为 $p_1, p_2, ..., p_{12}$，为 12 个近似独立的随机变量，其概率密度函数记为 $f_1, f_2, ..., f_{12}$，服从正态分布，记其均值分别为 $\mu_1, \mu_2, ..., \mu_{12}$，标准差分别为 $\sigma_1, \sigma_2, ..., \sigma_{12}$. 由于时间序列分析得到的未来一年的销量本身就反映了图书在过去的销售过程中体现出来的热度，因此我们可以利用这 12 个预测出的随机变量来进行热度分析. 为方便，我们将 2021 年 4 月记为第 1 个月，将 2021 年 5 月记为第 2 个月，以此类推，则 2022 年 3 月为第 12 个月.

记第 i 个月的库存为 w_i, 则：

$$w_i = l - k - (p_1 + p_2 + \cdots + p_{i-1}), i = 1, 2, .., 12.$$

显然，当第 i 个月的库存量无法满足当月的销量的时候便可以开始下一次印刷了，我们取

$$w_i - p_i < b \tag{15}$$

时开始下一次印刷，b 为一常数，印刷量为：

$$h = cw_i.$$

其中，c 为一变动系数. 为根据热度确定下一次的印刷量，我们设第 i 个月的热度指标为 $d_i, i = 1, 2, ..., 12$，它由第 i 个月到第 12 个月的销量决定，其公式为

$$d_i = \begin{cases} \frac{\mu_{i+1}+\mu_{i+2}+\cdots+\mu_{12}}{\mu_i+\mu_{i+1}+\cdots+\mu_{12}}, & i = 1, 2, ..., 11 \\ 0, & i = 12. \end{cases} \tag{16}$$

则 d_i 越大，表示第 $i+1$ 个月到最后一个月的总销量占第 i 个月到最后一个月的总销量的比重大，那么我们有理由认为相对第 i 月来说，未来一段时间的销量热度也比较大，那么它所蕴藏的潜在价值也越大，在目标函数中带来的收益就越应该被重视.

设第 i 个月进行印刷，印刷的成本 s 为

$$s = \begin{cases} 0.32rh, & h \geq 10000, \text{单色} \\ 0.34rh, & h < 10000, \text{单色} \\ 0.35rh, & h \geq 10000, \text{双色} \\ 0.41rh, & h < 10000, \text{双色} \\ 0.45rh, & h \geq 10000, \text{四色} \\ 0.54rh, & h < 10000, \text{四色} \end{cases} \tag{17}$$

则包含热度指标的利润可定义为：

$$\begin{aligned} y(p_i + p_{i+1} + \cdots + p_{12}) = &(1 + d_i)mn(p_i + p_{i+1} + \cdots + p_{12}) \\ &- s - 0.0273m(p_i + p_{i+1} + \cdots + p_{12}). \end{aligned} \tag{18}$$

由于假设了每月的销量近似独立，于是随机变量 $p_1 + p_2 + \cdots + p_{12}$ 服从正态分布 $N(\mu_1 + \mu_2 + \cdots + \mu_{12}, \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_{12}^2)$，记其概率密度函数为 $f^*(p_1 + p_2 + \cdots + p_{12})$.

于是就得到了可量化的包含了热度指标的利润：

$$y = E(y(p_1 + p_2 + \cdots + p_{12})) = \int_{-\infty}^{+\infty} y(p_1 + p_2 + \cdots + p_{12}) f^*(p_1 + p_2 + \cdots + p_{12}) \mathrm{d}x. \quad (19)$$

那么，我们的目标就是求使目标函数 (19) 最大化的系数 c.