1: Hybrid Gibbs Sampler to estimate Poisson Distribution λ (30%)

Derivation:

For these Poisson distributed random variables (r.v.s) ($n = 500$) with mean parameter λ, unobserved variables are $\lambda, y_1, y_2, \ldots, y_{78}$, in which $y's$ denote the r.v.s which are larger than or equal to five.

$$\begin{aligned}
&\text{Prior} : \pi(\lambda) \propto \frac{1}{\lambda};\\
&P(X, Y|\lambda) = \prod_{i=1}^{422} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \prod_{j=1}^{78} \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} I(y_i \geq 5);\\
&P(\lambda|X, Y) \propto P(X, Y|\lambda)P(\lambda) \propto e^{-n\lambda}\lambda^{\sum_i x_i + \sum_j y_j - 1}\\
&P(y_j|X, \lambda) = \frac{e^{-\lambda}\lambda^{y_i}}{y_i!} I(y_i \geq 5) \propto \frac{\lambda^{y_i}}{y_i!} I(y_i \geq 5)
\end{aligned}$$

Then we can know that $\lambda|X, Y \propto Gamma(\sum_i x_i + \sum_j y_j, n)$.

The MH-Step to sample 78 unobserved $y's$ is

$$y_j^* = \begin{cases} y_j^{(t)} - 1, & \text{with probability } \frac{1}{3};\\ y_j^{(t)}, & \text{with probability } \frac{1}{3};\\ y_j^{(t)} + 1, & \text{with probability } \frac{1}{3}. \end{cases}$$

$$r = \min\left\{ \frac{[\lambda^{(t+1)}]^{y_j^*}/y_j^*!}{[\lambda^{(t+1)}]^{y_j^{(t)}}/y_j^{(t)}!} I(y_j^* \geq 5), 1 \right\}$$

where r is the accept-reject ratio.

Result: $\hat{\lambda} = 2.803$.

2: Gibbs Sampler for Clustering (30%)

Derivation:

We use $\left\{X_{ij}\right\}_{i=1,2,3,\ldots,1000}^{j=1,2,3}$ to denote the datum of i-th sample in j-th dimension. Then using the same notation in the question, the complete-data likelihood function is:

$$\begin{aligned}
f(X, Z|\Pi, \Theta) &= \prod_{i=1}^{1000} \prod_{k=1}^{3} \Big[P(Z_i = k|\Pi, \Theta) P(X_{ij}, j = 1, 2, 3|Z_j, \Pi, \Theta) \Big]^{I(Z_j = k)}\\
&= \prod_{i=1}^{1000} \prod_{k=1}^{3} \Big[P(Z_i = k|\Pi) \prod_{j=1}^{3} P(X_{ij}|Z_j, \Theta) \Big]^{I(Z_j = k)}\\
&= \prod_{i=1}^{1000} \prod_{k=1}^{3} \Big[\pi_k \prod_{j=1}^{3} \binom{10j}{X_{ij}} \theta_{jk}^{X_{ij}} (1 - \theta_{jk})^{10j - X_{ij}} \Big]^{I(Z_j = k)}
\end{aligned}$$

Note that given $Z_i = k$, for each sample i, $X_{ij} \sim Bino(10j, \theta_{jk})$; $P(\theta_{jk}) \propto Beta(a, b)$ (Prior: $P(\theta_{jk}) \propto Beta(1,1) \propto 1$); and $(\pi_1, \pi_2, \pi_3) \sim Dirichlet(\alpha_1, \alpha_2, \alpha_3)$, we can derive by the following:

$$\begin{aligned}
P(\Pi, \Theta, Z|X) &\propto P(\Pi)P(\Theta)P(X, Z|\Pi, \Theta) \\
&\propto \prod_{k=1}^{3} \pi_k^{\alpha_k - 1} \prod_{j=1}^{3}\prod_{k=1}^{3} \theta_{jk}^{a-1}(1-\theta_{jk})^{b-1} \prod_{i=1}^{1000}\prod_{k=1}^{3} \Big[\pi_k \prod_{j=1}^{3} \binom{10j}{X_{ij}} \theta_{jk}^{X_{ij}}(1-\theta_{jk})^{10j-X_{ij}}\Big]^{I(Z_j=k)}
\end{aligned}$$

For π:

$$\begin{aligned}
f(\Pi|\Theta, Z) &\propto \prod_{k=1}^{3} \pi_k^{\alpha_k - 1} \prod_{i=1}^{1000}\prod_{k=1}^{3} \pi_k^{I(Z_i=k)} \\
&\propto \prod_{k=1}^{3} \pi_k^{\alpha_k - 1} \prod_{k=1}^{3} [\pi_k]^{\sum_{i=1}^{1000} I(Z_i=k)} \\
&\propto \prod_{k=1}^{3} \pi_k^{\alpha_k + \sum_{i=1}^{1000} I(Z_i=k) - 1}
\end{aligned}$$

Then it follows that:

$$\Pi|\Theta, Z \sim Dirichlet(\alpha_1 + \sum_{i=1}^{1000} I(Z_i = 1), \alpha_2 + \sum_{i=1}^{1000} I(Z_i = 2), \alpha_3 + \sum_{i=1}^{1000} I(Z_i = 3))$$

For θ:

$$\begin{aligned}
\theta_{jk}|- &\propto \prod_{j=1}^{3}\prod_{k=1}^{3} \theta_{jk}^{a-1}(1-\theta_{jk})^{b-1} \prod_{i=1}^{1000}\prod_{k=1}^{3} \Big[\prod_{j=1}^{3} \binom{10j}{X_{ij}} \theta_{jk}^{X_{ij}}(1-\theta_{jk})^{10j-X_{ij}}\Big]^{I(Z_j=k)} \\
&\propto \theta_{jk}^{a-1}(1-\theta_{jk})^{b-1} \prod_{i=1}^{1000} \Big[\theta_{jk}^{X_{ij}}(1-\theta_{jk})^{10j-X_{ij}}\Big]^{I(Z_j=k)} \\
&\propto \theta_{jk}^{\sum_{i=1}^{1000} X_{ij} I(Z_i=k) + a - 1}(1-\theta_{jk})^{\sum_{i=1}^{1000}(10j - X_{ij})I(Z_i=k) + b - 1}
\end{aligned}$$

Then it follows that:

$$\theta_{jk}|- \sim Beta(a + \sum_{i=1}^{1000} X_{ij} I(Z_i = k), b + \sum_{i=1}^{1000}(10j - X_{ij})I(Z_i = k))$$

For Z:

$$\begin{aligned}
Z_i|- &\propto \prod_{i=1}^{1000}\prod_{k=1}^{3} \pi_k^{I(Z_i=k)} \prod_{i=1}^{1000}\prod_{k=1}^{3}\prod_{j=1}^{3} \binom{10j}{X_{ij}}^{I(Z_i=k)} \theta_{jk}^{X_{ij} I(Z_i=k)} [(1-\theta_{jk})^{10j-X_{ij}}]^{I(Z_i=k)} \\
&\propto \prod_{k=1}^{3} \pi_k^{I(Z_i=k)} \prod_{k=1}^{3}\prod_{j=1}^{3} \binom{10j}{X_{ij}}^{I(Z_i=k)} \theta_{jk}^{X_{ij} I(Z_i=k)} [(1-\theta_{jk})^{10j-X_{ij}}]^{I(Z_i=k)} \\
&\propto \prod_{k=1}^{3} \Bigg[\pi_k \Big[\prod_{j=1}^{3} \binom{10j}{X_{ij}} \theta_{jk}^{X_{ij}}(1-\theta_{jk})^{10j-X_{ij}}\Big]\Bigg]^{I(Z_i=k)}
\end{aligned}$$

Then for Gibbs Sampler algorithm:

Given $\Pi^{(t)}, \Theta^{(t)}, Z^{(t)}$, we update the parameters by the following:

$$\pi_1^{(t+1)}, \pi_2^{(t+1)}, \pi_3^{(t+1)}|- \sim Dirichlet(\alpha_1 + \sum_{i=1}^{1000} I(Z_i^{(t)} = 1), \alpha_2 + \sum_{i=1}^{1000} I(Z_i^{(t)} = 2), \alpha_3 + \sum_{i=1}^{1000} I(Z_i^{(t)} = 3))$$

$$\theta_{jk}^{(t+1)}|- \sim Beta(a + \sum_{i=1}^{1000} X_{ij} I(Z_i^{(t)} = k), b + \sum_{i=1}^{1000} (10j - X_{ij}) I(Z_i^{(t)} = k))$$

$$P(Z_i^{(t+1)} = k|-) = \frac{\pi_k^{(t+1)} \prod_{j=1}^{3} \binom{10j}{X_{ij}} (\theta_{jk}^{(t+1)})^{X_{ij}} (1 - \theta_{jk}^{(t+1)})^{10j - X_{ij}}}{\sum_{l=1}^{3} \pi_l^{(t+1)} \prod_{j=1}^{3} \binom{10j}{X_{ij}} (\theta_{jl}^{(t+1)})^{X_{ij}} (1 - \theta_{jl}^{(t+1)})^{10j - X_{ij}}}$$

Result:

Estimation of Π:

$\hat{\pi_1} = 0.499$, $\hat{\pi_2} = 0.298$, and $\hat{\pi_3} = 0.203$.

Estimation of Θ:

$\theta_{11} = 0.807$, $\theta_{12} = 0.490$, $\theta_{13} = 0.196$, $\theta_{21} = 0.206$, $\theta_{22} = 0.804$, $\theta_{23} = 0.515$, $\theta_{31} = 0.502$, $\theta_{32} = 0.196$, and $\theta_{33} = 0.797$.

Estimation of Z:

Samples of cluster 1:

```
> which(estimated_z==1)
  [1]   2   4   8  12  14  15  18  19  20  21  23  24  26  27  28  32  33  36  38  39  40  44  45  47  48  49  50
 [28]  51  52  54  55  57  60  61  63  67  68  69  70  71  72  73  74  76  78  79  80  82  84  89  90  92  93  94
 [55]  96  99 104 106 108 112 113 115 118 120 122 123 128 129 130 131 135 137 140 141 143 146 147 150 152 153 154
 [82] 155 156 157 162 163 164 166 172 173 174 175 180 181 184 185 186 187 188 189 191 192 196 198 200 202 204 209
[109] 212 213 215 216 221 224 225 226 232 233 236 237 238 239 240 241 242 243 245 251 252 253 258 260 262 263 267
[136] 268 269 270 274 278 279 280 281 282 283 284 285 286 290 293 295 296 300 301 302 307 308 310 311 312 315 318
[163] 320 321 323 324 325 326 328 329 330 331 334 335 336 337 339 344 346 347 350 351 352 353 357 358 359 365 366
[190] 367 369 370 371 373 374 379 381 383 387 388 389 393 394 395 396 401 402 404 406 407 411 412 413 414 416 417
[217] 418 419 420 423 424 425 427 428 429 430 432 434 435 436 440 444 446 453 455 456 460 461 471 472 474 475 478
[244] 479 481 483 487 488 491 492 494 498 499 500 501 502 511 513 515 516 517 520 521 523 525 526 527 528 529 530
[271] 534 535 537 540 541 542 545 548 550 552 554 555 557 558 563 567 571 575 577 578 579 581 587 588 590 592 594
[298] 596 597 599 601 604 608 613 614 615 617 619 620 622 623 624 625 629 632 634 637 638 641 642 643 647 649 650
[325] 651 652 657 658 660 664 666 667 668 669 672 673 675 676 679 681 682 683 686 688 690 692 695 697 699 700 703
[352] 704 706 707 708 709 710 713 715 716 717 718 720 721 725 726 728 730 734 735 738 741 742 743 745 746 748 749
[379] 752 753 755 758 759 760 761 764 773 779 780 782 783 786 787 790 792 797 798 805 808 809 810 812 814 815 816
[406] 817 818 820 824 826 832 833 834 836 838 840 841 842 844 845 848 850 853 854 855 856 861 862 865 866 868 869
[433] 870 871 875 878 880 881 882 883 884 885 886 887 888 889 892 893 894 895 896 897 904 906 908 910 911 915 917
[460] 919 920 922 924 925 926 930 931 934 936 940 941 946 952 953 954 955 957 960 961 962 963 964 965 968 972 973
[487] 976 977 981 982 985 988 992 994 995 996 997 998 999
```

Figure 1: Samples in cluster 1

Samples of cluster 2:

```
> which(estimated_z==2)
  [1]   3   5   6   7  10  11  13  16  17  25  29  31  34  37  42  43  46  56  58  65  66
 [22]  75  81  88  91  97  98 100 105 107 109 114 116 119 121 124 126 127 133 134 136 139
 [43] 142 144 145 148 151 159 160 161 165 169 170 171 176 182 183 194 195 197 199 203 205
 [64] 206 210 217 219 220 223 227 229 230 231 234 235 247 249 254 264 271 272 273 275 276
 [85] 277 289 291 292 294 298 304 305 309 313 316 317 319 322 332 342 345 348 349 354 360
[106] 361 362 363 368 372 375 376 378 382 391 397 405 409 415 422 431 433 437 438 439 441
[127] 443 445 447 448 449 450 451 454 457 462 463 473 485 486 489 490 495 496 497 504 506
[148] 507 512 531 532 536 539 543 546 549 553 560 561 562 564 568 573 580 584 585 586 591
[169] 593 595 598 600 602 605 606 609 610 612 618 626 633 639 640 646 656 659 661 663 665
[190] 670 674 677 680 684 687 689 693 696 701 705 711 712 714 719 724 727 733 737 739 744
[211] 750 754 757 762 763 766 770 771 774 775 777 784 785 788 791 793 795 796 799 800 802
[232] 803 806 811 821 822 823 825 827 828 829 831 835 837 839 843 846 849 852 857 863 864
[253] 867 872 873 874 876 879 890 898 902 903 905 907 912 913 914 918 921 928 929 932 933
[274] 935 937 939 942 943 944 945 947 948 949 951 966 967 969 970 974 975 978 979 984 987
[295] 989 990 993 1000
```

Figure 2: Samples in cluster 2

Samples of cluster 3:

```
> which(estimated_z==3)
  [1]   1   9  22  30  35  41  53  59  62  64  77  83  85  86  87  95 101 102 103 110 111 117 125 132 138 149 158
 [28] 167 168 177 178 179 190 193 201 207 208 211 214 218 222 228 244 246 248 250 255 256 257 259 261 265 266 287
 [55] 288 297 299 303 306 314 327 333 338 340 341 343 355 356 364 377 380 384 385 386 390 392 398 399 400 403 408
 [82] 410 421 426 442 452 458 459 464 465 466 467 468 469 470 476 477 480 482 484 493 503 505 508 509 510 514 518
[109] 519 522 524 533 538 544 547 551 556 559 565 566 569 570 572 574 576 582 583 589 603 607 611 616 621 627 628
[136] 630 631 635 636 644 645 648 653 654 655 662 671 678 685 691 694 698 702 722 723 729 731 732 736 740 747 751
[163] 756 765 767 768 769 772 776 778 781 789 794 801 804 807 813 819 830 847 851 858 859 860 877 891 899 900 901
[190] 909 916 923 927 938 950 956 958 959 971 980 983 986 991
```

Figure 3: Samples in cluster 3

Traceplots:

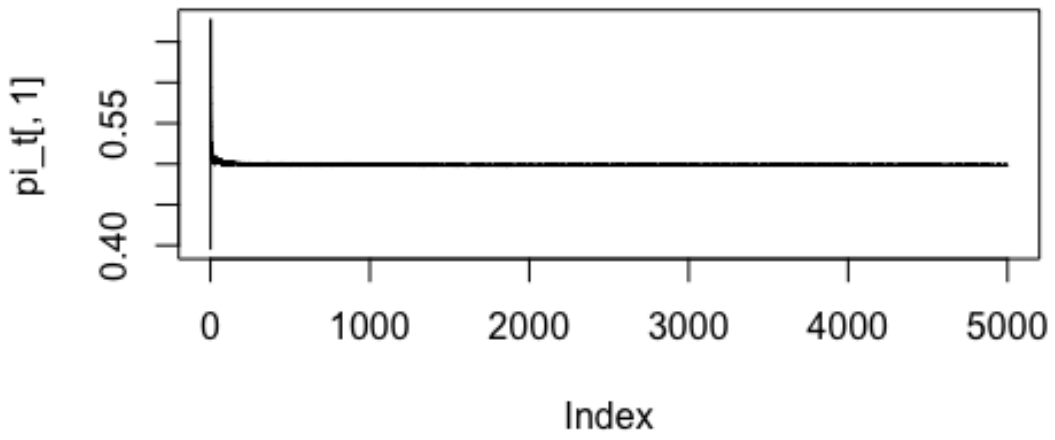

Figure 4: Traceplot of π_1

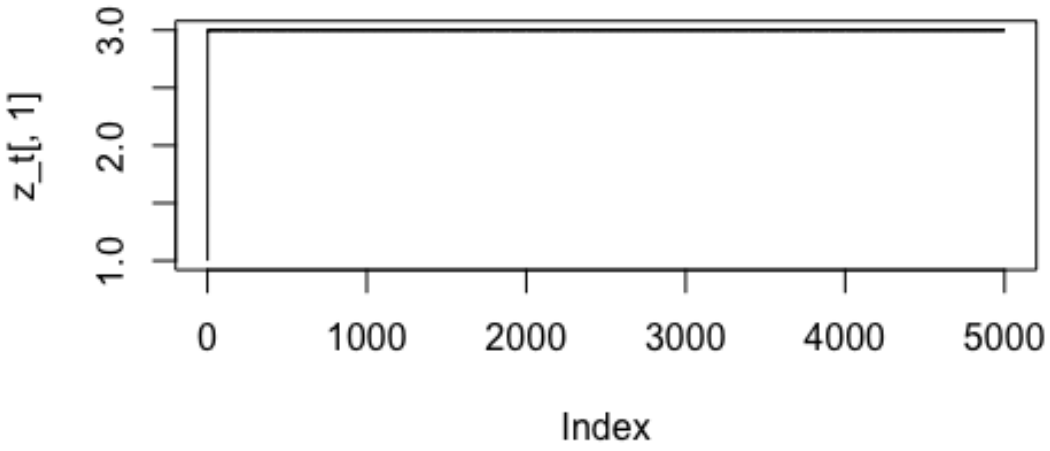

Figure 5: Traceplot of Z_1

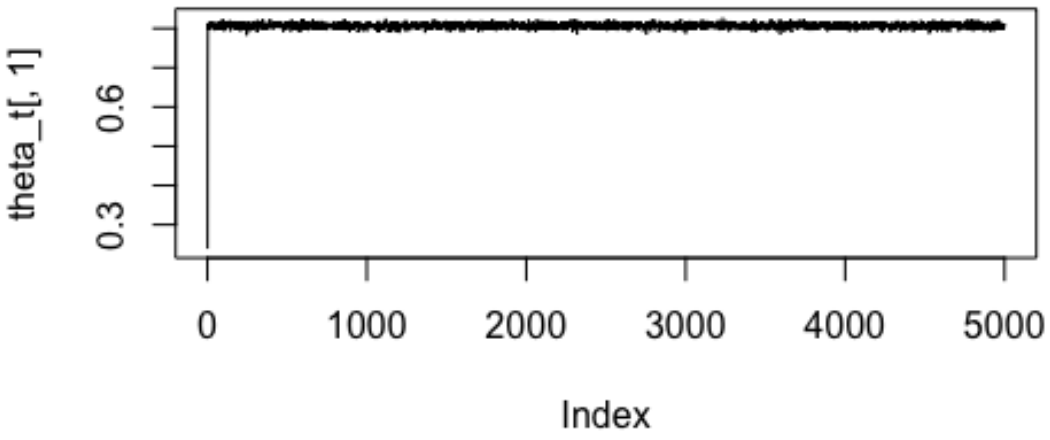

Figure 6: Traceplot of θ_{11}

3: Hybrid Gibbs Sampler (40%)

Note that for $i = 1, 2$, $\mathbf{Y}_i = (y_{i1}, y_{i2}, y_{i3}, y_{i4}) \sim multinomial(100, p_1, p_2, p_3, p_4)$, then:

$$\text{Prior}: \pi(\mathbf{p}) \propto Dirichlet(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \propto Dirichlet(2, 2, 2, 2) \propto p_1 p_2 p_3 p_4;$$

$$P(y_{i1}, y_{i2}, y_{i3}, y_{i4}|p_1, p_2, p_3, p_4) = \frac{100!}{y_{i1}!y_{i2}!y_{i3}!y_{i4}!} p_1^{y_{i1}} p_2^{y_{i2}} p_3^{y_{i3}} p_4^{y_{i4}}$$

$$\begin{aligned} P(p_1, p_2, p_3, p_4|y_{i1}, y_{i2}, y_{i3}, y_{i4}) &\propto P(y_{i1}, y_{i2}, y_{i3}, y_{i4}|p_1, p_2, p_3, p_4) f(p_1, p_2, p_3, p_4|\alpha_1, \alpha_2, \alpha_3, \alpha_4) \\ &\propto \frac{100!}{y_{i1}!y_{i2}!y_{i3}!y_{i4}!} p_1^{y_{i1}+\alpha_1-1} p_2^{y_{i2}+\alpha_2-1} p_3^{y_{i3}+\alpha_3-1} p_4^{y_{i4}+\alpha_4-1} \end{aligned}$$

$$p_1, p_2, p_3, p_4|y_{i1}, y_{i2}, y_{i3}, y_{i4} \sim Dirichlet(y_{i1} + \alpha_1, y_{i2} + \alpha_2, y_{i3} + \alpha_3, y_{i4} + \alpha_4)$$

$$P(y_{12}|\mathbf{Y}, \mathbf{P}) \propto \frac{p_2^{y_{12}}}{y_{12}!} \frac{p_1^{y_{11}}}{y_{11}!} = \frac{p_2^{y_{12}}}{y_{12}!} \frac{p_1^{47-y_{12}}}{(47-y_{12})!}$$

$$P(y_{22}|\mathbf{Y}, \mathbf{P}) \propto \frac{p_2^{y_{22}}}{y_{22}!} \frac{p_4^{y_{24}}}{y_{24}!} = \frac{p_2^{y_{22}}}{y_{22}!} \frac{p_4^{46-y_{22}}}{(46-y_{22})!}$$

Then MH-Step to update y_{i2} (for $i = 1, 2$) is

$$y_{i2}^{(t+1)} = \begin{cases} y_{i2}^{(t)} + 1, & \text{with probability } \frac{1}{2} \\ y_{i2}^{(t)} - 1, & \text{with probability } \frac{1}{2} \end{cases}$$

$$y_{i2}^{(t+1)} = y_{i2}^{(t)} + 1, \qquad \text{if } y_{i2}^{(t)} = 15$$

$$y_{i2}^{(t+1)} = y_{i2}^{(t)} - 1, \qquad \text{if } y_{i2}^{(t)} = 32$$

$$r = \begin{cases} \min\left\{2 \times \frac{P(y_{i2}^{(t+1)}|\mathbf{Y},\mathbf{P})}{P(y_{i2}^{(t)}|\mathbf{Y},\mathbf{P})}, 1\right\}, & \text{if } y_{i2}^{(t+1)} = 32 \text{ or } 15 \\ \min\left\{\frac{1}{2} \times \frac{P(y_{i2}^{(t+1)}|\mathbf{Y},\mathbf{P})}{P(y_{i2}^{(t)}|\mathbf{Y},\mathbf{P})}, 1\right\}, & \text{if } y_{i2}^{(t)} = 32 \text{ or } 15 \\ \min\left\{\frac{P(y_{i2}^{(t+1)}|\mathbf{Y},\mathbf{P})}{P(y_{i2}^{(t)}|\mathbf{Y},\mathbf{P})}, 1\right\}, & \text{otherwise} \end{cases}$$

where r is the accept-reject ratio. Then for another two unobserved variables:

$$y_{11}^{(t+1)} = 100 - y_{13} - y_{14} - y_{12}^{(t+1)} = 100 - 22 - 31 - y_{12}^{(t+1)}$$

$$y_{24}^{(t+1)} = 100 - y_{21} - y_{23} - y_{22}^{(t+1)} = 100 - 28 - 26 - y_{22}^{(t+1)}$$

Result:

$p_1 = 0.298$, $p_2 = 0.153$, $p_3 = 0.241$, and $p_4 = 0.308$.